

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-1
Sequence: 1 NYHDIYLAAGVCGSAKLVHQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_21.*
2: geneseqp1908.*
3: geneseqp1908.*
4: geneseqp2001.*
5: geneseqp2001.*
6: geneseqp2003.*
7: geneseqp2003.*
8: geneseqp2004.*
9: geneseqp2005.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	100.0	23	ADBI6886	AdBI6886 Haemophil
2	117	100.0	23	AEBO8218	AEBO8218 Haemophil
3	117	100.0	205	AAW50230	AAW50230 Haemophil
4	117	100.0	205	AEBO8261	AEBO8261 Haemophil
5	114	97.4	23	ADBI6912	AdBI6912 Haemophil
6	114	97.4	23	AEBO8246	AEBO8246 Haemophil
7	114	97.4	29	ADBI6916	AdBI6916 Escherich
8	114	97.4	29	AEBO8250	AEBO8250 Penetrati
9	95	81.2	23	ADBI6888	AdBI6888 Escherich
10	95	81.2	23	AEBO8220	AEBO8220 Escherich
11	95	81.2	24	ADBI6908	AdBI6908 Escherich
12	95	81.2	24	AEBO8242	AEBO8242 Escherich
13	95	81.2	25	ADBI6922	AdBI6922 Escherich
14	95	81.2	25	ADBI6906	AdBI6906 Escherich
15	95	81.2	26	ADBI6904	AdBI6904 E coli pe
16	95	81.2	29	ADBI6921	AdBI6921 E coli pe
17	95	81.2	30	ADBI6903	AdBI6903 E coli pe
18	95	81.2	30	ADBI6905	AdBI6905 Escherich
19	95	81.2	30	AEBO8239	AEBO8239 Penetrati
20	95	81.2	31	ADBI6918	AdBI6918 Escherich
21	95	81.2	31	AEBO8252	AEBO8252 Penetrati
22	95	81.2	213	AEBO8263	AEBO8263 Escherich
23	93	79.5	217	AEBO61272	AEBO61272 Klebsiell
24	92	78.6	23	ADBI6887	AdBI6887 Pasteurel

25	92	78.6	23	9	AEBO8219	AEBO8219 Pasteurel
26	92	78.6	203	9	AEBO8262	AEBO8262 Pasteurel
27	89	76.1	23	6	ADBI6909	AdBI6909 Escherich
28	89	76.1	23	9	AEBO8243	AEBO8243 Escherich
29	89	76.1	30	6	ADBI6919	AdBI6919 Escherich
30	89	76.1	30	6	AEBO8253	AEBO8253 Penetrati
31	82	70.1	22	6	ADBI6910	AdBI6910 Escherich
32	82	70.1	22	6	AEBO8244	AEBO8244 Escherich
33	82	70.1	29	6	ADBI6920	AdBI6920 Escherich
34	82	70.1	29	6	AEBO8254	AEBO8254 Penetrati
35	78	66.7	209	6	AEBO67069	AEBO67069 Phototrab
36	76	65.0	216	7	ADFO7708	ADFO7708 Bacterial
37	72	61.5	23	6	ADBI6990	AdBI6990 Buchnera
38	72	61.5	23	9	AEBO8222	AEBO8222 Buchnera
39	72	61.5	211	9	AEBO8265	AEBO8265 Buchnera
40	56	47.9	23	6	ADBI6889	AdBI6889 Vibrio ch
41	56	47.9	23	9	AEBO8221	AEBO8221 Vibrio ch
42	56	47.9	204	9	AEBO8264	AEBO8264 Vibrio ch
43	48	41.0	23	6	ADBI6892	AdBI6892 Xylella f
44	48	41.0	23	9	AEBO8224	AEBO8224 Xylella f
45	48	41.0	204	9	AEBO8267	AEBO8267 Xylella f

ALIGNMENTS

RESULT 1
ADBI6886
ID ADBI6886 standard; peptide; 23 AA.

ADBI6886;

20-NOV-2003 (first entry)

Haemophilus influenzae ORF HI0638 penetrating peptide 1.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;
infertility; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiant; antitartaroclastic;
osteopathic; cytostatic; nootropic.

Haemophilus influenzae.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-IB000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating
or preventing e.g. endocrine disorders.

Claim 2, Page 14; 60pp; English.

This invention relates to a novel peptide sequences capable of
translocating across a biological barrier. Furthermore, it refers to
methods that use these peptides to facilitate penetration of a
biologically active effector molecule such as a drug or other therapeutic
agent across biological barriers e.g. epithelial or endothelial cells
sealed by tight junctions. This peptide is derived from a bacterial
toxin, an integral membrane or extracellular protein and can comprise an
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
or enzyme. The effector molecule, however, can comprise for example
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptide of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from ORF HI0638 of
 CC Haemophilus influenzae and is penetrating peptide 1 of the invention.
 XX

Sequence 23 AA:

Query Match 100.0%; Score 117; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.9e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIVLAVGVCOSAKLVHQLA 23
 |||||
 Db 1 NYHDIVLAVGVCOSAKLVHQLA 23

RESULT 2
 AEB08218
 ID AEB08218 standard; peptide; 23 AA.
 XX
 AC AEB08218;
 DT 25-AUG-2005 (first entry)
 DX
 DE Haemophilus influenzae ORF HI0638 penetrating peptide 1, SEQ ID NO: 1.
 XX
 KM Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KM antidiabetic; endocrine disease; gastrointestinal disease;
 KM metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KM degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KM neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KM neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KM dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KM anticonvulsant; genetic disorder; cardiovascular disease;
 KM cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KM coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KM nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KM genitourinary disease; hematological disease; antianemic; anemia;
 KM autoimmune disease; immunosuppressive; immune deficiency;
 KM immunosuppressant; infectious disease; antimicrobial; infection;
 KM erectile dysfunction; andrology; major depressive disorder;
 KM antidepressant; psychiatric disorder; pain; analgesic;
 KM bacterial infection; antibacterial; viral infection; virocidic;
 KM fungal infection; fungicide; parasitic infection; antiparasitic;
 KM renal failure; antifertility; antirheumatic; cyostatic;
 KM antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Haemophilus influenzae.
 XX
 PN US2005136103-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 16-SEP-2004; 2004US-00942300.
 XX
 PR 17-SEP-2003; 2003US-00664989.
 XX
 PR 17-SEP-2003; 2003US-00665184.
 XX
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 XX
 PA (COHEN/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 DR WPI, 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 46; SEQ ID NO 1; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transcellular delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Haemophilus influenzae ORF
 CC HI0638 penetrating peptide. This sequence is used in the effective
 CC translocation of aminoglycoside antibiotics and antifungal agents across
 CC an epithelial barrier.
 XX

Sequence 23 AA:

Query Match 100.0%; Score 117; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.9e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIVLAVGVCOSAKLVHQLA 23
 |||||
 Db 1 NYHDIVLAVGVCOSAKLVHQLA 23

RESULT 3
 AAMS0230
 ID AAMS0230 standard; protein; 205 AA.
 XX
 AC AAMS0230;
 XX
 DT 07-JAN-2002 (first entry)
 DX
 DE Haemophilus influenzae protein HI 0638, involved in paracytosis.
 XX
 KM ORF HI 0638; paracytosis; paracytin; lung; epithelium;
 KM lower respiratory tract; infection; antimicrobial; drug screening;
 KM therapy; vaccine.
 XX
 OS Haemophilus influenzae.
 XX
 PN EP1136557-A1.
 XX
 PD 26-SEP-2001.
 XX
 PF 21-MAR-2000; 2000EP-00201045.
 XX
 PR 21-MAR-2000; 2000EP-00201045.
 XX
 PA (NEWE-) NEDERLANDEN MIN WELZIJN.
 XX
 PA Van Den Schifgaarde, Van Alphen AJW;
 XX
 PI WPI, 2001-640390/74.
 XX
 DR N-PSDB; AAI70333.
 XX
 PT New nucleic acids and polypeptides, which increases the permeability of
 XX tissue or cell layer, useful as vaccine for treating lower respiratory

tract infections due to Haemophilus influenzae.

Claim 5, Page 19, 32pp; English.

The present sequence is that of protein HI 0638, which is encoded by newly identified open reading frame (ORF) HI 0638 (see AAI70333) of Haemophilus influenzae strain A960053, an isolate from a sputum sample of a cystic fibrosis patient. The protein is involved in paracytosis. Clones of B. coli DHS- α containing ORF HI 0636 (see AAI70332) and/or ORF HI 0638 showed significant increase in penetration of epithelial cell layers of the human bronchial epithelial cell line NCI-H292. HI 0636 (see AAM50229) and HI 0638 may be used to alter the permeability of tissues or cell layers, especially lung epithelial cell layers. They will also be useful for developing compounds that block paracytosis as well as to obtain insight into the invasion mechanism used by various bacteria. This may lead to novel compounds and/or approaches in the treatment of lower respiratory tract infections due to H. influenzae. Vaccines containing or based upon either the HI 0636 and/or HI 0638 protein or polypeptide, or a microorganism expressing ORF HI 0636 and/or ORF HI 0638, are claimed

Sequence 205' AA;

Query Match 100.0%; Score 117; DB 4; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NYHDIYALAGVCSAKLVHOLA 23
3 NYHDIYALAGVCSAKLVHOLA 25

RESULT 4
AEB08261
ID AEB08261 standard; protein; 205 AA.

AEB08261;
25-AUG-2005 (first entry)

Haemophilus influenzae ORF HI0638 protein, SEQ ID NO: 59.

pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes; antidiabetic; endocrine disease; gastrointestinal disease; metabolic disorder; hormone deficiency; osteoporosis; osteopathic; degeneration; musculoskeletal disease; ocular disease; neurodegenerative disease; neuroprotective; Alzheimer's disease; neurotropic; neurological disease; parkinson's disease; antiparkinsonian; dementia; multiple sclerosis; immune disorder; Huntington's chorea; cardiovascular-gen.; atherosclerosis; cardiovascular disease; anticonvulsant; genetic disorder; cardiovascular disease; coronary artery disease; cardiac; vasodilator; obesity; anorectic; nutritional disorder; vitamin deficiency; renal disease; nephrotoxic; genitourinary disease; hematological disease; anti-anemic; anemia; autoimmune disease; immunosuppressive; immune deficiency; immunostimulant; infectious disease; antimicrobial; infection; erectile dysfunction; andrology; major depressive disorder; antidepressant; psychiatric disorder; pain; analgesic; bacterial infection; antibacterial; viral infection; virolytic; fungal infection; fungicide; parasitic infection; antiparasitic; renal failure; antifertility; antirheumatic; cytostatic; antiinflammatory; hepatotropic; hepatitis B virus infection.

Haemophilus influenzae.

US2005136103-A1.

23-JUN-2005.

16-SEP-2004; 2004US-00942300.

17-SEP-2003; 2003US-00664989.

17-SEP-2003; 2003US-00665184.

17-SEP-2003; 2003US-0503615P.

(BENS/) BEN-SASSON S A.
(COHE/) COHEN E.

Ben-Sasson SA, Cohen E;
WPI: 2005-444089/45.

Composition used for translocating effector across barrier such as epithelial cells during treatment of e.g. endocrine disorders comprises effector sequentially coupled with counter ion and hydrophobic agent.

Disclosure; SEQ ID NO 59; 59pp; English.

The present invention relates to a pharmaceutical composition of penetrating peptides for transmembrane delivery of effector. The invention comprises the effector sequentially coupled with a counter ion and at least one hydrophobic agent, where the effector is selectively encapsulated into a complex. The invention is useful for translocating effectors across a biological barrier such as epithelial cells and endothelial cells during treatment and prevention of disease or pathological conditions (including endocrine disorders, diabetes, infertility, hormone deficiencies, osteoporosis, ophthalmological disorders, neurodegenerative disorders, Alzheimer's disease, dementia, Parkinson's disease, multiple sclerosis, Huntington's disease, cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypocoagulable states, coronary disease, cerebrovascular events, metabolic disorders, obesity, vitamin deficiencies, renal disorders, renal failure, hematological disorders, anemia of different entities, immunologic and rheumatologic disorders, autoimmune diseases, immune deficiencies, infectious diseases, viral infections, bacterial infections, fungal infections, parasitic infections, neoplastic diseases, multi-factorial disorders, impotence, chronic pain, depression, different fibrosis states and short stature) and for mucosal vaccination against anthrax and hepatitis B. The present sequence is the Haemophilus influenzae ORF HI0638 protein containing penetrating peptide at N-terminal end.

Sequence 205' AA;

Query Match 100.0%; Score 117; DB 9; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NYHDIYALAGVCSAKLVHOLA 23
3 NYHDIYALAGVCSAKLVHOLA 25

RESULT 5
ADBI6912
ID ADBI6912 standard; peptide; 23 AA.

ADBI6912;
20-NOV-2003 (first entry)

Haemophilus influenzae ORF HI0638 penetrating peptide 34.

penetrating peptide; epithelial; endothelial; tight junction; diabetes; infertility; hormone; vitamin deficiency; neurodegenerative; cardiovascular; haematological; endocrine disorder; obesity; neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic; osteopathic; cytostatic; neurotropic.

Haemophilus influenzae.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-IB000968.

07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA Ben-Sasson SA, Cohen E;
 XX WPI, 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PS Claim 2, Page 15, 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antihypertensive, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from ORF HI0638 of
 CC Haemophilus influenzae and is penetrating peptide 34 of the invention.
 SQ Sequence 23 AA;
 QY
 Db 1 NYHDIVALAGVCSAKLVHQLA 23
 1 NYHDIVALAGVCSAKLVHQLA 23
 RESULT 6
 AEB08246 97.4%; Score 114; DB 6; Length 23;
 ID AEB08246 standard; peptide; 23 AA.
 AC AEB08246;
 XX
 XX 25-AUG-2005 (first entry)
 DT
 XX
 DE Haemophilus influenzae ORF HI0638 penetrating peptide 34, SEQ ID NO: 29.
 XX
 XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; arteriosclerosis;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antipneumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.

XX
 OS Haemophilus influenzae.
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI, 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 PS Claim 46; SEQ ID NO 29; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Haemophilus influenzae ORF
 CC HI0638 penetrating peptide. This sequence is used in the effective
 CC translocation of aminoglycoside antibiotics and antifungal agents across
 CC an epithelial barrier.
 SQ Sequence 23 AA;
 QY
 Db 1 NYHDIVALAGVCSAKLVHQLA 23
 1 NYHDIVALAGVCSAKLVHQLA 23
 RESULT 7
 ADB16916 97.4%; Score 114; DB 9; Length 23;
 ID ADB16916 standard; peptide; 29 AA.
 AC ADB16916;
 XX
 XX 20-NOV-2003 (first entry)
 DT
 XX
 DE Escherichia coli peptide 3 coupled to imaging compound linker, IBW-006.
 XX
 XX epithelial; endothelial; tight junction; diabetes; diabetes; infertility; hormone;

vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 endocrine disorder; obesity; neoplastic disease; neuroprotective;
 cartilage; antiarteriosclerotic; osteopathic; cytostatic; neurotrophic;
 imaging linker; penetrating peptide; IBM-006.
 06 Synthetic.
 06 Escherichia coli.
 Key Location/Qualifiers
 Modified-site 1 /note="N-terminal acetyl"
 Modified-site 29 /note="C-terminal amide"
 MO2003066859-A2.
 14-AUG-2003.
 07-FEB-2003; 2003WO-IB000968.
 07-FEB-2002; 2002US-035396P.
 (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 Ben-Sasson SA, Cohen E;
 WPI; 2003-697452/66.
 New penetrating peptide, useful for preparing a composition for treating
 or preventing e.g. endocrine disorders.
 Example 3; Page 40; 60pp; English.
 This invention relates to a novel peptide sequences capable of
 translocating across a biological barrier. Furthermore, it refers to
 methods that use these peptides to facilitate penetration of a
 biologically active effector molecule such as a drug or other therapeutic
 agent across biological barriers e.g. epithelial or endothelial cells
 sealed by tight junctions. This peptide is derived from a bacterial
 toxin, an integral membrane or extracellular protein and can comprise an
 anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 or enzyme. The effector molecule, however, can comprise for example
 insulin, gonadotropin, erythropoietin, granulocyte/macrophage colony
 stimulating factor (GM-CSF), enkephalin, dalaargin, or neurotrophic
 factors. The penetrating peptide is useful for the treatment of various
 conditions including diabetes, infertility, hormone and vitamin
 deficiencies, neurodegenerative, cardiovascular, haematological and
 endocrine disorders, as well as obesity and neoplastic disease.
 Accordingly, the peptides of this invention can be used in compositions
 that have neuroprotective, cartilage, antiarteriosclerotic, osteopathic,
 cytostatic or neurotrophic activities. This peptide sequence is IBM-006,
 which consists of the Escherichia coli penetrating peptide 3 coupled to
 the imaging linker peptide used in an exemplification of the invention.
 Sequence 29 AA;
 SQ
 Query Match 97.4%; Score 114; DB 6; Length 29;
 Best Local Similarity 95.7%; Pred. No. 1.1e-10;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYHDIYVLAGVCSASLVHQLA 23
 1 NYHDIYVLAGVCSASLVHQLA 23
 DB 1 NYHDIYVLAGVCSASLVHQLA 23
 RESULT 8
 AEB08250 ID AEB08250 standard; peptide; 29 AA.
 AC AEB08250;
 XX 25-AUG-2005 (first entry)
 DT XX

Penetrating peptide SEQ. 33 used in composition for mucosal vaccination.
 06 Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 antidiabetic; endocrine disease; gastrointestinal disease;
 metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 neurodegenerative disease; neuroprotective; Alzheimer's disease;
 neurotrophic; neurological disease; parkinson's disease; antiparkinsonian;
 dementia; multiple sclerosis; immune disorder; Huntington chorea;
 anticonvulsant; genetic disorder; cardiovascular disease;
 cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 coronary artery disease; cardiomyopathy; obesity; anorectic;
 nutritional disorder; vitamin deficiency; renal disease; nephroretropic;
 genitourinary disease; hematological disease; antineutic; anemia;
 autoimmune disease; immunosuppressive; immune deficiency;
 immunostimulant; infectious disease; antimicrobial; infection;
 erectile dysfunction; andrology; major depressive disorder;
 antidepressant; psychiatric disorder; pain; analgesic;
 bacterial infection; antibacterial; viral infection; virocidic;
 fungal infection; fungicide; parasitic infection; antiparasitic;
 renal failure; antifertility; antineumatic; cytostatic;
 antiinflammatory; hepatotropic; hepatitis B virus infection.
 Unidentified.
 Key Location/Qualifiers
 Misc-difference 1 /note="N-terminal acylated"
 Misc-difference 29 /note="Optionally C-terminal amide, optionally the free
 amino group of lysine is acylated with a fatty acid"
 US2005136103-A1.
 23-JUN-2005.
 16-SEP-2004; 2004US-00942300.
 17-SEP-2003; 2003US-00664989.
 17-SEP-2003; 2003US-00665184.
 17-SEP-2003; 2003US-0503615P.
 (BENS/) BEN-SASSON S A.
 (COHE/) COHEN E.
 Ben-Sasson SA, Cohen E;
 WPI; 2005-444089/45.
 Composition used for translocating effectors across barrier such as
 epithelial cells during treatment of e.g. endocrine disorders comprises
 effector sequentially coupled with counter ion and hydrophobic agent.
 Claim 63; SEQ ID NO 33; 59pp; English.
 The present invention relates to a pharmaceutical composition of
 penetrating peptides for transmembrane delivery of effector. The
 invention comprises the effector sequentially coupled with a counter ion
 and at least one hydrophobic agent, where the effector is selectively
 encapsulated into a complex. The invention is useful for translocating
 effectors across a biological barrier such as epithelial cells and
 endothelial cells during treatment and prevention of disease or
 pathological conditions (including endocrine disorders, diabetes,
 infertility, hormone deficiencies, osteoporosis, ophthalmological
 disorder, neurodegenerative disorders, Alzheimer's disease, dementia,
 Parkinson's disease, multiple sclerosis, Huntington's disease,
 cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 coagulable states, coronary disease, cerebrovascular events, metabolic
 disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 hematological disorders, anemia of different entities, immunologic and
 rheumatologic disorders, autoimmune diseases, immune deficiencies,
 infectious diseases, viral infections, bacterial infections, fungal
 infections, parasitic infections, neoplastic diseases, multi-factorial

CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is a penetrating peptide (IBM-006) used
 CC in the composition for mucosal vaccination using a counter anion and a
 CC penetrating peptide.

SQ Sequence 29 AA;

Query Match 97.4%; Score 114; DB 9; Length 29;
 Best Local Similarity 95.7%; Pred. No. 1.1e-10;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYHDIYVLAGVCSAKLVHQLA 23
 |||||
 DB 1 NYHDIYVLAGVCSAKLVHQLA 23

RESULT 9
 ADB16888
 ID ADB16888 standard; peptide; 23 AA.
 XX
 AC ADB16888;

DT 20-NOV-2003 (first entry)

DE Escherichia coli YCFC penetrating peptide 3.

KM penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 KM infertility; hormone; vitamin deficiency; neurodegenerative;
 KM cardiovascular; haematological; endocrine disorder; obesity;
 KM neoplastic disease; neuroprotective; cardiac; arteriosclerotic;
 KM osteopathic; cyostatic; nootropic.

XX Escherichia coli.

OS Escherichia coli.

PN WO2003066859-A2.

PD 14-AUG-2003.

PP 07-FEB-2003; 2003WO-1B000968.

PR 07-FEB-2002; 2002US-0355396P.

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PI Ben-Sasson SA, Cohen E;

DR WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.

PS Claim 2, Page 14; 60pp; English.

CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factor. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiatherosclerotic, osteopathic,
 CC cyostatic or nootropic activities. This peptide is from YCFC of
 CC Escherichia coli and is penetrating peptide 3 of the invention.

XX SQ Sequence 23 AA;

Query Match 81.2%; Score 95; DB 6; Length 23;
 Best Local Similarity 78.3%; Pred. No. 9.1e-08;
 Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NYHDIYVLAGVCSAKLVHQLA 23
 |||||
 DB 1 NYHDIYVLAGVCSAKLVHQLA 23

RESULT 10
 AEB08220
 ID AEB08220 standard; peptide; 23 AA.
 XX
 AC AEB08220;

DT 25-AUG-2005 (first entry)

DE Escherichia coli YCFC penetrating peptide 3, SEQ ID NO: 3.

KM Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KM antidiabetic; endocrine disease; gastrointestinal disease;
 KM metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KM degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KM neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KM nootropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KM dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KM anticonvulsant; genetic disorder; cardiovascular disease;
 KM cardiovascular-gen.; atherosclerosis; arteriosclerotic;
 KM coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KM nutritional disorder; vitamin deficiency; renal disease; nephrologic;
 KM genitourinary disease; hematological disease; anemic; anemia;
 KM autoimmune disease; immunosuppressive; immune deficiency;
 KM immunostimulant; infectious disease; antimicrobial; infection;
 KM erectile dysfunction; andrology; major depressive disorder;
 KM antidepressant; psychiatric disorder; pain; analgesic;
 KM bacterial infection; antibacterial; viral infection; virocidic;
 KM fungal infection; fungicide; parasitic infection; antiparasitic;
 KM renal failure; antifertility; antineumatic; cyostatic;
 KM antiinflammatory; hepatotropic; hepatitis B virus infection.

XX Escherichia coli.

OS US2005136103-A1.

PN 23-JUN-2005.

PD 16-SEP-2004; 2004US-00942300.

PP 17-SEP-2003; 2003US-00664989.

PR 17-SEP-2003; 2003US-00665184.

PR 17-SEP-2003; 2003US-0503615P.

PA (BENS// BEN-SASSON S A.

PI (COHEN// COHEN E.

DR WPI; 2005-444089/45.

PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.

PS Claim 53; SEQ ID NO 3; 59pp; English.

CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating

effectors across a biological barrier such as epithelial cells and endothelial cells during treatment and prevention of disease or pathological conditions (including endocrine disorders, diabetes, infertility, hormone deficiencies, osteoporosis, ophthalmological disorders, neurodegenerative disorders, Huntington's disease, dementia, Parkinson's disease, multiple sclerosis, hyper-coagulable states, hypocoagulable states, coronary disease, cerebrovascular events, metabolic disorders, obesity, vitamin deficiencies, renal disorders, renal failure, hematological disorders, anemia of different entities, immunologic and rheumatologic disorders, autoimmune diseases, immune deficiencies, infectious diseases, viral infections, bacterial infections, fungal infections, parasitic infections, neoplastic diseases, multi-factorial disorders, impotence, chronic pain, depression, different fibrosis states and short stature) and for mucosal vaccination against anthrax and hepatitis B. The present sequence is the Escherichia coli YCFC penetrating peptide. This sequence is used in the effective translocation of aminoglycoside antibiotics and antifungal agents across an epithelial barrier.

Sequence 23 AA;

Query Match 81.2%; Score 95; DB 9; Length 23;
Best Local Similarity 78.3%; Pred. No. 9.1e-08;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCGSARLVQQLA 23
DB 1 NYHDIYALAGVCGSARLVQQLA 23

RESULT 11

ADB16908 standard; peptide; 24 AA.

ADB16908;

20-NOV-2003 (first entry)

Escherichia coli YCFC penetrating peptide 30.

penetrating peptide; epithelial; endothelial; tight junction; diabetes; infertility; hormone; vitamin deficiency; neurodegenerative; cardiovascular; haematological; endocrine disorder; obesity; neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic; osteopathic; cytosolic; nootropic.

Escherichia coli.

MO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003MO-IB000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating or preventing e.g. endocrine disorders.

Claim 2; Page 15; 60pp; English.

This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells

sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, hematological and endocrine disorders, as well as obesity and neoplastic disease. Accordingly, the peptides of this invention can be used in compositions that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic, cytosolic or nootropic activities. This peptide is from YCFC of Escherichia coli and is penetrating peptide 30 of the invention.

Sequence 24 AA;

Query Match 81.2%; Score 95; DB 6; Length 24;
Best Local Similarity 78.3%; Pred. No. 9.6e-08;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCGSARLVQQLA 23
DB 2 NYHDIYALAGVCGSARLVQQLA 24

RESULT 12

AE08242 standard; peptide; 24 AA.

AE08242;

25-AUG-2005 (first entry)

Escherichia coli YCFC penetrating peptide 30, SEQ ID NO: 25.

pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes; antidiabetic; endocrine disease; gastrointestinal disease; metabolic disorder; hormone deficiency; osteoporosis; osteopathic; degeneration; musculoskeletal disease; ocular disease; ophthalmological; neurodegenerative disease; neuroprotective; Alzheimer's disease; nootropic; neurological disease; Parkinson's disease; antiparkinsonian; dementia; multiple sclerosis; immune disorder; Huntingtons chorea; cardiovascular; genetic disorder; cardiovascular disease; cardiovascular-gen.; atherosclerosis; antiarteriosclerotic; coronary artery disease; cardiac; vasotropic; obesity; nephrotropic; nutritional disorder; vitamin deficiency; renal disease; nephrotropic; genitourinary disease; hematological disease; anemia; autoimmune disease; immunosuppressive; immune deficiency; immunostimulant; infectious disease; antimicrobial; infection; erectile dysfunction; andrology; major depressive disorder; antidepressant; psychiatric disorder; pain; analgesic; bacterial infection; antibacterial; viral infection; viruside; fungal infection; fungicide; parasitic infection; antiparasitic; renal failure; antifertility; antirheumatic; cytosolic; antiinflammatory; hepatotropic; hepatitis B virus infection.

Escherichia coli.

US2005136103-A1.

23-JUN-2005.

16-SEP-2004; 2004US-00942300.

17-SEP-2003; 2003US-00664989.

17-SEP-2003; 2003US-00665184.

17-SEP-2003; 2003US-0503615P.

(BENS/) BEN-SASSON S A.

(COHE/) COHEN E.

Ben-Sasson SA, Cohen E;

XX WPI; 2005-444089/45.
 DR Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 PS Claim 46; SEQ ID NO 25; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatological disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YCC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 XX
 SQ Sequence 24 AA;
 Query Match 81.2%; Score 95; DB 9; Length 24;
 Best Local Similarity 78.3%; Pred. No. 9.6e-08;
 Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY
 1 NYHDIVTLAGVCSAKLVHQLA 23
 ||:|||||:|||||:|||||
 Db 2 NYDITTLAAGICQSAKRLVQQLA 24
 RESULT 13
 ADB16922 standard; peptide; 25 AA.
 XX ADB16922;
 AC
 XX 20-NOV-2003 (first entry)
 DT
 XX Escherichia coli peptide 3 linked to recombinant human insulin.
 DE
 XX recombinant human insulin; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cytosstatic; nootropic; penetrating peptide; diabetes.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "Penetrating peptide 3"
 FT Modified-site 25
 FT /note= "Recombinant human insulin peptide coupled to the
 FT C-terminus"
 XX WO2003066859-A2.
 XX

PD 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 PF
 XX 07-FEB-2002; 2002US-0355396P.
 PR
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA
 XX Ben-Saason SA, Cohen E;
 PI WPI; 2003-697452/66.
 DR
 XX
 XX
 XX
 PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PS Example 8; Page 42; 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cytosstatic or nootropic activities. This peptide sequence consists of the
 CC Escherichia coli penetrating peptide 3 coupled to recombinant human
 CC insulin in the absence of a detachable linker peptide as a cleavage site,
 CC used to deliver insulin across mucosal epithelia as a treatment for
 CC diabetes, in an exemplification of the invention.
 XX
 XX
 SQ Sequence 25 AA;
 Query Match 81.2%; Score 95; DB 6; Length 25;
 Best Local Similarity 78.3%; Pred. No. 1e-07;
 Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY
 1 NYHDIVTLAGVCSAKLVHQLA 23
 ||:|||||:|||||:|||||
 Db 1 NYDITTLAAGICQSAKRLVQQLA 23
 RESULT 14
 ADB16906 standard; peptide; 25 AA.
 XX ADB16906;
 AC
 XX 20-NOV-2003 (first entry)
 DT
 XX Escherichia coli peptide 3 linked to linearised insulin receptor peptide.
 DE
 XX linearised insulin receptor; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cytosstatic; nootropic; penetrating peptide; diabetes.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "Penetrating peptide 3"
 FT

```

FT Modified-site 25
FT /note="Linearised insulin receptor peptide coupled to
FT the C-terminus"
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-1B000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX WPI, 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Example 9; Page 43; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells
XX sealed by tight junctions. This peptide is derived from a bacterial
XX toxin, an integral membrane or extracellular protein and can comprise an
XX antioesulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX or enzyme. The effector molecule, however, can comprise for example
XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
XX factors. The penetrating peptide is useful for the treatment of various
XX conditions including diabetes, infertility, hormone and vitamin
XX deficiencies, neurodegenerative, cardiovascular, haematological and
XX endocrine disorders, as well as obesity and neoplastic disease.
XX Accordingly, the peptides of this invention can be used in compositions
XX that have neuroprotective, cardiant, antihypertensive, osteopathic,
XX cyostatic or neurotropic activities. This peptide sequence consists of the
XX Escherichia coli penetrating peptide 3 coupled to a linearised insulin
XX receptor, used to deliver insulin across mucosal epithelia as a treatment
XX for diabetes, in an exemplification of the invention.
XX
XX Sequence 25 AA;
XX
XX Query Match 81.2%; Score 95; DB 6; Length 25;
XX Best Local Similarity 78.3%; Pred. No. 1e-07;
XX Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 NYHDIVLALAGVGSASLVQQLA 23
XX ||:|||||:|||||:|||||
XX 1 NYDITLALAGVGSASLVQQLA 23
XX
XX RESULT 15
XX ADB16904
XX ID ADB16904 standard; peptide; 26 AA.
XX
XX AC ADB16904;
XX
XX DT 20-NOV-2003 (first entry)
XX
XX DE E_coli penetrating peptide 3 linked to heparin without a cleavage site.
XX
XX KM heparin; epithelial; endothelial; tight junction; diabetes; infertility;
XX hormone; vitamin deficiency; neurodegenerative; cardiovascular;
XX haematological; endocrine disorder; obesity; neoplastic disease;
XX KM neuroprotective; cardiant; antihypertensive; osteopathic; cyostatic;
XX neurotropic; penetrating peptide.
XX
XX OS Synthetic.

```

```

OS Escherichia coli.
XX
XX Key Location/Qualifiers
XX Peptide 1..23
XX
XX Modified-site 26
XX /note="Penetrating peptide 3"
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-1B000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX WPI, 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Example 10; Page 44; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells
XX sealed by tight junctions. This peptide is derived from a bacterial
XX toxin, an integral membrane or extracellular protein and can comprise an
XX antioesulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX or enzyme. The effector molecule, however, can comprise for example
XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
XX factors. The penetrating peptide is useful for the treatment of various
XX conditions including diabetes, infertility, hormone and vitamin
XX deficiencies, neurodegenerative, cardiovascular, haematological and
XX endocrine disorders, as well as obesity and neoplastic disease.
XX Accordingly, the peptides of this invention can be used in compositions
XX that have neuroprotective, cardiant, antihypertensive, osteopathic,
XX cyostatic or neurotropic activities. This peptide sequence consists of the
XX Escherichia coli penetrating peptide 3 coupled to heparin in the absence
XX of a detachable linker peptide as a cleavage site, used to deliver low
XX molecular weight heparin across mucosal epithelia in an exemplification
XX of the invention.
XX
XX Sequence 26 AA;
XX
XX Query Match 81.2%; Score 95; DB 6; Length 26;
XX Best Local Similarity 78.3%; Pred. No. 1e-07;
XX Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 NYHDIVLALAGVGSASLVQQLA 23
XX ||:|||||:|||||:|||||
XX 1 NYDITLALAGVGSASLVQQLA 23
XX
XX Search completed: January 23, 2006, 09:20:46
XX Job time : 70.2 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:56 ; Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838A-1

Perfect score: 117

Sequence: 1 NYHDIVALAGVCSAKLVHQLA 23

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications AA.New:*
1: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pep:*
5: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pep:*
7: /cgn2_6/prodata/2/pubppa/US11_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	35.9	42	6	US-10-512-295A-3
2	40	34.2	307	6	US-10-793-626-684
3	39.5	33.8	447	6	US-10-858-730-219
4	38.5	32.9	305	6	US-10-520-820-1
5	38.5	32.9	365	7	US-11-000-463-241
6	38.5	32.9	365	7	US-11-000-463-713
7	37.5	32.1	447	7	US-11-055-822-94
8	37	31.6	306	7	US-11-017-550-45
9	37	31.6	306	7	US-11-014-402-5
10	37	31.6	306	7	US-11-137-850-6
11	37	31.6	388	7	US-11-046-668-7
12	37	31.6	398	7	US-11-046-668-9
13	37	31.6	412	6	US-10-979-821-8
14	37	31.6	412	7	US-11-114-922-8
15	37	31.6	2644	6	US-10-770-726-45
16	36	30.8	141	7	US-11-074-176-104
17	36	30.8	147	7	US-11-074-176-14
18	36	30.8	204	6	US-10-858-730-95
19	36	30.8	231	7	US-11-082-389-242
20	36	30.8	250	7	US-11-051-670-2
21	36	30.8	322	6	US-10-689-742-46
22	36	30.8	359	7	US-11-194-246-293
23	36	30.8	502	6	US-10-063-703-158
24	36	30.8	502	7	US-11-102-240-158
25	36	30.8	513	7	US-11-055-822-348

26	36	30.8	787	6	US-10-467-657-2832	Sequence 2832, App
27	36	30.8	810	6	US-10-453-372-1116	Sequence 1116, App
28	36	30.8	1027	6	US-10-793-626-3106	Sequence 3106, App
29	36	30.8	2144	7	US-11-124-368A-277	Sequence 277, App
30	35.5	30.3	162	6	US-10-467-657-2510	Sequence 2510, App
31	35.5	30.3	463	6	US-10-531-844-2	Sequence 2, App1
32	35.5	30.3	730	6	US-10-821-234-1019	Sequence 1019, App
33	35	29.9	52	6	US-10-516-768-28	Sequence 28, App1
34	35	29.9	52	6	US-10-516-768-29	Sequence 29, App1
35	35	29.9	52	6	US-10-529-118-3	Sequence 3, App1
36	35	29.9	113	6	US-10-467-657-2448	Sequence 2448, App
37	35	29.9	120	6	US-10-793-626-2404	Sequence 2404, App
38	35	29.9	164	6	US-10-467-657-6308	Sequence 6308, App
39	35	29.9	180	5	US-09-978-360A-463	Sequence 463, App
40	35	29.9	185	6	US-10-821-234-1498	Sequence 1498, App
41	35	29.9	185	6	US-10-529-118-2	Sequence 2, App1
42	35	29.9	185	6	US-10-453-372-896	Sequence 896, App
43	35	29.9	194	6	US-10-453-372-894	Sequence 894, App
44	35	29.9	205	6	US-10-793-626-1828	Sequence 1828, App
45	35	29.9	218	6	US-10-453-372-892	Sequence 892, App

ALIGNMENTS

RESULT 1
US-10-512-295A-3
; Sequence 3, Application US/10512295A
; Publication No. US20050245727A1
GENERAL INFORMATION:
; APPLICANT: Baltzer, Lars
; APPLICANT: Dolphin, Gunnar
; APPLICANT: Liedberg, Bo
; APPLICANT: Lundstrom, Ingemar
; TITLE OF INVENTION: NOVEL POLYPEPTIDE SCAFFOLDS AND USE THEREOF
; FILE REFERENCE: 5848.18USWO
; CURRENT APPLICATION NUMBER: US/10/512,295A
; PRIOR FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/SE03/00507
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: SE 0200968-6
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide KE3
US-10-512-295A-3
Query Match 35.9%; Score 42; DB 6; Length 42;
Best Local Similarity 52.9%; Pred. No. 1.2;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Cy 7 LALAGVCSAKLVHQLA 23
Db 16 LALAGVCSAKLVHQLA 32
RESULT 2
US-10-793-626-684
; Sequence 684, Application US/10793626
; Publication No. US20050255478A1
GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09


```

US-11-000-463-713
; Sequence 713, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radcoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIPACN
; CURRENT FILING DATE: 2004-11-29
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 713
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-713

Query Match      32.9%; Score 38.5; DB 7; Length 365;
Best Local Similarity 43.5%; Pred. No. 46;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY      4 DIVLALAGVC---QSATLVHQLA 23
DB      67 DLMGLAGFCGSMASGHLFKQMA 89

RESULT 7
US-11-055-822-94
; Sequence 94, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Schroder, Burthard
; APPLICANT: Schneider, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberman, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12

```

```

; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 94
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-94

Query Match      32.1%; Score 37.5; DB 7; Length 447;
Best Local Similarity 33.3%; Pred. No. 83;
Matches 10; Conservative 6; Mismatches 3; Indels 11; Gaps 1;

QY      1 NYHDIVL-----ALAGVCSAKIV 19
DB      9 NYDMILKRNAGEPEFHQAIVAEVLSLKV 38

RESULT 8
US-11-017-550-45
; Sequence 45, Application US/11017550
; Publication No. US20050250183A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Schultz, Peter G
; APPLICANT: Wang, Lei
; APPLICANT: Anderson, John C
; APPLICANT: Chin, Jason
; APPLICANT: Liu, David R
; APPLICANT: Magliery, Thomas
; APPLICANT: Meggers, Eric L
; APPLICANT: Mehl, Ryan A
; APPLICANT: Pasternak, Mitro
; APPLICANT: Santoro, Stephen W
; APPLICANT: Zhang, Zhiwen
; TITLE OF INVENTION: In Vivo Incorporation of Unnatural Amino Acids
; FILE REFERENCE: 54-000120US
; CURRENT APPLICATION NUMBER: US/11/017,550
; CURRENT FILING DATE: 2004-12-17
; PRIOR APPLICATION NUMBER: US/10/126,927
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,030
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 60/355,514
; PRIOR FILING DATE: 2002-02-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 45
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-11-017-550-45

Query Match      31.6%; Score 37; DB 7; Length 306;
Best Local Similarity 34.8%; Pred. No. 67;
Matches 8; Conservative 7; Mismatches 6; Indels 2; Gaps 1;

QY      1 NYHDIVLALAGVCSAKIVHQLA 23
DB      160 HHYGVDAVAGVGMGQ--RKIHMLA 180

```

```
RESULT 9
US-11-014-402-5
; Sequence 5, Application US/11014402
; Publication No. US20050266526A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Schultz, Peter
; APPLICANT: Wang, Lei
; APPLICANT: Zhang, Zhiwen
; TITLE OF INVENTION: GLYCOPROTEIN SYNTHESIS
; FILE REFERENCE: 54A-000610US
; CURRENT APPLICATION NUMBER: US/11/014,402
; CURRENT FILING DATE: 2004-12-14
; PRIOR APPLICATION NUMBER: US/10/686,944
; PRIOR FILING DATE: 2003-10-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: mutant synthetase derived from Methanococcus jannaschii cytosyl-t
US-11-014-402-5

Query Match      31.6%; Score 37; DB 7; Length 306;
Best Local Similarity 34.8%; Pred. No. 67;
Matches 8; Conservative 7; Mismatches 6; Indels 2; Gaps 1;

QY      1 NYHDIYLAGVQSAKLVHOLA 23
Db      160 HTHGVAVGMEQ--RKIHMLA 180

RESULT 10
US-11-137-850-6
; Sequence 6, Application US/11137850
; Publication No. US20050272121A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: IRM, LLC
; APPLICANT: Xie, Jianming
; APPLICANT: Wang, Lei
; APPLICANT: Wu, Ning
; APPLICANT: Schultz, Peter G
; APPLICANT: Sprengon, Glen
; TITLE OF INVENTION: SITE SPECIFIC INCORPORATION OF HEAVY ATOM-CONTAINING UNNATURAL
; FILE REFERENCE: 54-000920US
; CURRENT APPLICATION NUMBER: US/11/137,850
; CURRENT FILING DATE: 2005-05-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: mutant synthetase
US-11-137-850-6

Query Match      31.6%; Score 37; DB 7; Length 306;
Best Local Similarity 34.8%; Pred. No. 67;
Matches 8; Conservative 7; Mismatches 6; Indels 2; Gaps 1;

QY      1 NYHDIYLAGVQSAKLVHOLA 23
Db      160 HTHGVAVGMEQ--RKIHMLA 180

RESULT 11
US-11-046-668-7
```

```
; Sequence 7, Application US/11046668
; Publication No. US20050244855A1
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: Novel Human Kinase Proteins and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0227-USA
; CURRENT APPLICATION NUMBER: US/11/046,668
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US/09/940,921
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/229,280
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 388
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-046-668-7

Query Match      31.6%; Score 37; DB 7; Length 388;
Best Local Similarity 26.3%; Pred. No. 86;
Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY      4 DIVYLAGVQSAKLVHOL 22
Db      203 DTILFMKQICGIRHMQM 221

RESULT 12
US-11-046-668-9
; Sequence 9, Application US/11046668
; Publication No. US20050244855A1
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: Novel Human Kinase Proteins and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0227-USA
; CURRENT APPLICATION NUMBER: US/11/046,668
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US/09/940,921
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/229,280
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 398
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-046-668-9

Query Match      31.6%; Score 37; DB 7; Length 398;
Best Local Similarity 26.3%; Pred. No. 88;
Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY      4 DIVYLAGVQSAKLVHOL 22
Db      203 DTILFMKQICGIRHMQM 221

RESULT 13
US-10-979-821-8
; Sequence 8, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAMOW, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
US-10-979-821-8
```

APPLICANT: HICKS, PAULA M.
APPLICANT: MCPARLAN, SARA C.
APPLICANT: MILLIS, JIM
APPLICANT: ROSAZZA, JACK
APPLICANT: ZHAO, LISHAN
APPLICANT: WEINER, DAVID P.
TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
FILE REFERENCE: 023829-0390
CURRENT APPLICATION NUMBER: US/10/979,821
CURRENT FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: 10/422,366
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 60/374,831
PRIOR FILING DATE: 2002-04-23
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentin Ver. 3.3
SEQ ID NO 8
LENGTH: 412
TYPE: PRT
ORGANISM: Leishmania major
US-10-979-821-8

Query Match 31.6%; Score 37; DB 6; Length 412;
Best Local Similarity 33.3%; Pred. No. 92;
Matches 7; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HDIVLALAGVCSAKLVHOLA 23
DB 375 HNIPTVSGRANMAGLTHERA 395

RESULT 14
US-11-114-922-8
Sequence 8, Application US/11114922
Publication No. US20050282260A1
GENERAL INFORMATION:
APPLICANT: HICKS, PAULA M.
APPLICANT: MCPARLAN, SARA C.
TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
FILE REFERENCE: 023829-0396
CURRENT APPLICATION NUMBER: US/11/114,922
CURRENT FILING DATE: 2005-04-26
PRIOR APPLICATION NUMBER: 10/422,366
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 60/374,831
PRIOR FILING DATE: 2002-04-23
NUMBER OF SEQ ID NOS: 91
SOFTWARE: Patentin Ver. 3.3
SEQ ID NO 8
LENGTH: 412
TYPE: PRT
ORGANISM: Leishmania major
US-11-114-922-8

Query Match 31.6%; Score 37; DB 7; Length 412;
Best Local Similarity 33.3%; Pred. No. 92;
Matches 7; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HDIVLALAGVCSAKLVHOLA 23
DB 375 HNIPTVSGRANMAGLTHERA 395

RESULT 15
US-10-770-726-45
Sequence 45, Application US/10770726
Publication No. US20050266409A1
GENERAL INFORMATION:
APPLICANT: Wyeth
APPLICANT: Brown, Eugene
APPLICANT: Liu, Wei

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
TITLE OF INVENTION: CANCERS
FILE REFERENCE: AM101079 (031896-010000)
CURRENT APPLICATION NUMBER: US/10/770,726
CURRENT FILING DATE: 2004-02-04
NUMBER OF SEQ ID NOS: 48640
SOFTWARE: Patentin version 3.2
SEQ ID NO 45
LENGTH: 2644
TYPE: PRT
ORGANISM: Homo sapiens
US-10-770-726-45

Query Match 31.6%; Score 37; DB 6; Length 2644;
Best Local Similarity 27.3%; Pred. No. 6.7e+02;
Matches 6; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 1 NYHDIVLALAGVCSAKLVHOL 22
DB 1735 HYHGVKSMGLGQLSTVITQV 1756

Search completed: January 23, 2006, 11:53:35
Job time : 6.1 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 seconds

(without alignments)
114.551 Million cell updates/sec

Title: US-10-501-838A-1

Perfect score: 117

Sequence: 1 NYHDIYALAGVCGSAKLVIHQQLA 23

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5.COMB.pep:*
2: /cgn2_6/prodata/1/aa/6.COMB.pep:*
3: /cgn2_6/prodata/1/aa/H.COMB.pep:*
4: /cgn2_6/prodata/1/aa/PCTUS.COMB.pep:*
5: /cgn2_6/prodata/1/aa/RE.COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	79.5	217	US-09-489-039A-7789	Sequence 7789, Ap
2	76	65.0	216	US-09-543-681A-7993	Sequence 7993, Ap
3	52.5	44.9	367	US-09-404-296B-6	Sequence 6, Appl
4	47	40.2	309	US-09-902-540-12828	Sequence 12828, A
5	46	39.3	138	US-09-270-767-58510	Sequence 58510, A
6	46	39.3	265	US-09-328-352-4347	Sequence 4347, Ap
7	46	39.3	433	US-09-270-767-43172	Sequence 43172, A
8	44	37.6	398	US-09-543-681A-7046	Sequence 7046, Ap
9	42.5	36.3	158	US-08-828-832-3	Sequence 3, Appl
10	42.5	36.3	159	US-09-496-632C-16	Sequence 16, Appl
11	42.5	36.3	159	US-08-828-832-1	Sequence 1, Appl
12	42.5	36.3	159	US-09-496-632C-15	Sequence 15, Appl
13	42	35.9	121	US-09-270-767-39953	Sequence 39953, A
14	42	35.9	121	US-09-270-767-55170	Sequence 55170, A
15	42	35.9	246	US-09-199-637A-407	Sequence 407, App
16	42	35.9	272	US-09-813-453B-5	Sequence 5, Appl
17	42	35.9	272	US-09-712-363-276	Sequence 276, App
18	42	35.9	313	US-09-270-767-33519	Sequence 33519, A
19	42	35.9	470	US-09-769-863-20	Sequence 20, Appl
20	41	35.0	287	US-09-252-991A-29808	Sequence 29808, A
21	41	35.0	303	US-09-543-681A-7924	Sequence 7924, Ap
22	41	35.0	436	US-09-605-703B-2864	Sequence 2864, Ap
23	41	35.0	471	US-09-252-991A-27633	Sequence 27633, A
24	40.5	34.6	427	US-09-248-796A-17203	Sequence 17203, A
25	40.5	34.6	717	US-09-107-532A-6350	Sequence 6350, Ap
26	40	34.2	90	US-09-124-671-25	Sequence 25, Appl
27	40	34.2	109	US-09-124-671-23	Sequence 23, Appl

28	40	34.2	109	2	US-09-124-671-27	Sequence 27, Appl
29	40	34.2	109	2	US-09-124-671-29	Sequence 29, Appl
30	40	34.2	110	1	US-08-244-537-10	Sequence 10, Appl
31	40	34.2	119	2	US-10-104-047-2840	Sequence 2840, Ap
32	40	34.2	165	2	US-09-134-000C-3704	Sequence 3704, Ap
33	40	34.2	181	2	US-09-134-001C-5665	Sequence 5665, Ap
34	40	34.2	183	2	US-09-134-000C-4230	Sequence 4230, Ap
35	40	34.2	233	2	US-09-489-039A-9070	Sequence 9070, Ap
36	40	34.2	272	2	US-09-489-039A-7644	Sequence 7644, Ap
37	40	34.2	307	2	US-09-710-279-684	Sequence 684, App
38	40	34.2	367	2	US-09-198-452A-316	Sequence 316, App
39	40	34.2	367	2	US-09-438-185A-304	Sequence 304, App
40	40	34.2	369	1	US-07-937-609-19	Sequence 19, Appl
41	40	34.2	369	2	US-08-029-170-19	Sequence 19, Appl
42	40	34.2	369	2	US-09-443-745-19	Sequence 19, Appl
43	40	34.2	420	1	US-08-846-762-73	Sequence 73, Appl
44	40	34.2	462	1	US-08-471-033-20	Sequence 20, Appl
45	40	34.2	462	1	US-08-471-044-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-489-039A-7789
Sequence 7789, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7789
LENGTH: 217
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match 79.5%; Score 93; DB 2; Length 217;
Best Local Similarity 78.3%; Pred. No. 3.5e-07;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

CY 1 NYHDIYALAGVCGSAKLVIHQQLA 23
DB 8 NYHDIYALAGVCGSAKLVIHQQLA 30

RESULT 2
US-09-543-681A-7993
Sequence 7993, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7993
LENGTH: 216
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7993

Query Match 65.0%; Score 76; DB 2; Length 216;

Best Local Similarity 56.5%; Pred. No. 0.00019;
Matches 13; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSAKLVHQA 23
Db 11 DFRDITLALAGICQASRLVQQA 33

RESULT 3

US-09-404-296B-6
; Sequence 6, Application US/09404296B
; Patent No. 6559358
; GENERAL INFORMATION:
; APPLICANT: MURRAY, James Augustus Henry
; TITLE OF INVENTION: PLANTS WITH MODIFIED GROWTH
; FILE REFERENCE: 2121-0151P
; CURRENT APPLICATION NUMBER: US/09/404,296B
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-404-296B-6

Query Match 44.9%; Score 52.5; DB 2; Length 367;
Best Local Similarity 45.8%; Pred. No. 2.2;
Matches 11; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

Qy 3 HDIVLAG-----VCQSAKLVHQ 21
Db 280 HDLIELMGTSGVNICQSLKRHQ 303

RESULT 4

US-09-902-540-12828
; Sequence 12828; Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12828
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12828

Query Match 40.2%; Score 47; DB 2; Length 309;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 8 ALAGVCSAKLVHQA 23
Db 152 ALAGVQEARLPHSPA 167

RESULT 5

US-09-270-767-58510
; Sequence 58510, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58510
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-58510

Query Match 39.3%; Score 46; DB 2; Length 138;
Best Local Similarity 52.9%; Pred. No. 8.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSAK 17
Db 24 NYADIELALAGIADKAR 40

RESULT 6

US-09-328-352-4347
; Sequence 4347, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4347
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4347

Query Match 39.3%; Score 46; DB 2; Length 265;
Best Local Similarity 58.8%; Pred. No. 17;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 LALAGVCSAKLVHQA 23
Db 41 LALAGVQATQLTMTA 57

RESULT 7

US-09-270-767-43172
; Sequence 43172, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43172
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURES:
; OTHER INFORMATION: Kaa means any amino acid
US-09-270-767-43172

Query Match 39.3%; Score 46; DB 2; Length 433;
Best Local Similarity 52.9%; Pred. No. 30;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSAK 17
|||:|||||:

Db 319 NYADIEIALAGLADKAR 335

RESULT 8
US-09-543-681A-7046

Sequence 7046, Application US/09543681A

Patent No. 6605709

GENERAL INFORMATION:

APPLICANT: GARY BRETON

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.1002-001

CURRENT APPLICATION NUMBER: US/09/543,681A

CURRENT FILING DATE: 2000-04-05

PRIOR APPLICATION NUMBER: US 60/128,706

PRIOR FILING DATE: 1999-04-09

NUMBER OF SEQ ID NOS: 8344

SEQ ID NO 7046

LENGTH: 398

TYPE: PRT

ORGANISM: Proteus mirabilis

US-09-543-681A-7046

Query Match 37.6%; Score 44; DB 2; Length 398;

Best Local Similarity 34.5%; Pred. No. 57;

Matches 10; Conservative 4; Mismatches 5; Indels 10; Gaps 1;

3 HDIVIALAG-----VCOSAKLVHQ 21

253 HDALVLRGKIKLSIRMEKICQNAQLIAQ 281

Db

RESULT 9

US-08-828-832-3

Sequence 3, Application US/08828832

Patent No. 5827711

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Shah, Puri

TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSER: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: IBM Compatible

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/828,832

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: pr-0250 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-845-4166

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 158 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 1575011

US-08-828-832-3

Query Match 36.3%; Score 42.5; DB 1; Length 158;

Best Local Similarity 31.8%; Pred. No. 36;

Matches 7; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

1 NYHDIVIALAGVCOSAKLVHQL 22

142 NYHDV-----GICKAVAMLMKL 158

Db

RESULT 10

US-09-496-632C-16

Sequence 16, Application US/09496632C

Patent No. 6468789

GENERAL INFORMATION:

APPLICANT: BAYSAI, Bora E.

APPLICANT: FERRELL, Robert E.

APPLICANT: DEVLIN, Bernie J.

APPLICANT: WILLETT-BROZICK, Joan E.

TITLE OF INVENTION: OXYGEN SENSING AND HYPOXIC SELECTION FOR TUMORS

FILE REFERENCE: 99-484-US

CURRENT APPLICATION NUMBER: US/09/496,632C

CURRENT FILING DATE: 2000-02-02

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn version 3.1

SEQ ID NO 16

LENGTH: 158

TYPE: PRT

ORGANISM: Bos taurus

US-09-496-632C-16

Query Match 36.3%; Score 42.5; DB 2; Length 158;

Best Local Similarity 31.8%; Pred. No. 36;

Matches 7; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

1 NYHDIVIALAGVCOSAKLVHQL 22

142 NYHDV-----GICKAVAMLMKL 158

Db

RESULT 11

US-08-828-832-1

Sequence 1, Application US/08828832

Patent No. 5827711

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Shah, Puri

TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSER: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: IBM Compatible

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/828,832

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0250 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 2454416
US-08-828-832-1

Query Match 36.3%; Score 42.5; DB 1; Length 159;
Best Local Similarity 31.8%; Pred. No. 36;
Matches 7; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

QY 1 NYHDIYALAGVCQSAKLVHQL 22
| | | | | : | | : | : | : | : |
DB 143 NYHDV-----GICKAVAMLMKL 159

RESULT 12
US-09-496-632C-15
Sequence 15, Application US/09496632C
Patent No. 6468769
GENERAL INFORMATION:
APPLICANT: BAYSAL, Bora E.
APPLICANT: FERRELL, Robert E.
APPLICANT: DEVLIN, Bernie J.
APPLICANT: WILLETT-BROZICK, Joan E.
TITLE OF INVENTION: OXYGEN SENSING AND HYPOXIC SELECTION FOR TUMORS
FILE REFERENCE: 99-484-US
CURRENT APPLICATION NUMBER: US/09/496,632C
CURRENT FILING DATE: 2000-02-02
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 159
TYPE: PRT
ORGANISM: Homo sapiens
US-09-496-632C-15

Query Match 36.3%; Score 42.5; DB 2; Length 159;
Best Local Similarity 31.8%; Pred. No. 36;
Matches 7; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

QY 1 NYHDIYALAGVCQSAKLVHQL 22
| | | | | : | | : | : | : | : |
DB 143 NYHDV-----GICKAVAMLMKL 159

RESULT 13
US-09-270-767-39953
Sequence 39953, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 39953
LENGTH: 121
TYPE: PRT

ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39953

Query Match 35.9%; Score 42; DB 2; Length 121;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 10; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 3 HDIVYALAG--VCQSAKLVH 20
| : | | | | | | | | | | | | | | | |
DB 1 HELVPLLVGVCVCQWQKRFVH 20

RESULT 14
US-09-270-767-55170
Sequence 55170, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 55170
LENGTH: 121
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-55170

Query Match 35.9%; Score 42; DB 2; Length 121;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 10; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 3 HDIVYALAG--VCQSAKLVH 20
| : | | | | | | | | | | | | | | | |
DB 1 HELVPLLVGVCVCQWQKRFVH 20

RESULT 15
US-09-199-637A-407
Sequence 407, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Mahajan-Miklos, Shailina
APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenthard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 407
LENGTH: 246
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-407

Query Match 35.9%; Score 42; DB 2; Length 246;
Best Local Similarity 42.9%; Pred. No. 70;

Matches	6;	Conservative	4;	Mismatches	4;	Indels	0;	Gaps	0;
Qy	2	YHDIYLAAGVQS	15						
	:	:	:						
Db	41	HHDIAGPAGICLS	54						

Search completed: January 23, 2006, 09:35:43
Job time : 16.6 secs

This Page Blank (usps.gov)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)

243.185 Million cell updates/sec

Title: US-10-501-838a-1

Sequence: 1 NYHDIYALAGVCSAKLVHQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	100.0	205	2 I64155	hypothetical prote
2	95	81.2	208	2 AH0139	conserved hypochet
3	95	81.2	213	2 S19211	ytic protein - Bsc
4	95	81.2	213	2 D90829	hypothetical prote
5	95	81.2	213	2 B85687	hypothetical prote
6	92	78.6	215	2 AB0647	conserved hypochet
7	72	61.5	211	2 D84960	hypothetical prote
8	56	47.9	205	2 B82237	conserved hypochet
9	48	41.0	204	2 E82680	conserved hypochet
10	47.5	40.6	531	2 C95338	hypothetical prote
11	46	39.3	947	2 AH0891	adenyl-transferrase
12	46	39.3	1265	2 T21782	hypothetical prote
13	46	39.3	1857	2 T50573	hypothetical prote
14	45.5	38.9	380	2 B86773	hypothetical prote
15	45.5	38.9	1215	2 T23916	malate dehydrogena
16	45	38.5	433	2 T25946	hypothetical prote
17	45	38.5	1243	2 T17390	hypothetical prote
18	45	38.5	2946	2 T00867	hypothetical prote
19	44	37.6	72	2 T66623	hypothetical prote
20	44	37.6	429	2 S09852	hypothetical prote
21	44	37.6	464	2 F64970	hypothetical prote
22	44	37.6	464	2 D50985	hypothetical prote
23	44	37.6	464	2 G85830	hypothetical prote
24	43	36.8	317	2 A44156	lipopolysaccharide
25	43	36.8	335	2 S70671	hypothetical prote
26	43	36.8	429	2 T24922	hypothetical prote
27	42	35.9	159	1 Q6ADE	hypothetical prote
28	42	35.9	159	1 Q6ADE	hypothetical prote
29	42	35.9	208	2 A98225	hypothetical prote

30	42	35.9	208	2 AF3061	ribonuclease D [Im
31	42	35.9	238	2 AH2178	orotidine 5' monop
32	42	35.9	272	2 A70955	hypothetical prote
33	42	35.9	342	2 AB0867	galactose operon r
34	42	35.9	511	2 T40334	hypothetical prote
35	42	35.9	517	2 T50801	argininosuccinate
36	42	35.9	1187	2 T18355	hypothetical prote
37	41.5	35.5	261	2 AC0164	probable transport
38	41	35.0	183	2 S24980	hydroxymethylbilan
39	41	35.0	193	2 B95340	hypothetical prote
40	41	35.0	261	2 B96036	probable phenylace
41	41	35.0	273	2 B83551	hypothetical prote
42	41	35.0	300	2 D83399	malate dehydrogena
43	41	35.0	415	2 AG0301	probable membrane
44	41	35.0	455	2 F82345	conserved hypochet
45	41	35.0	517	2 AC2237	uroporphyrinogen I

ALIGNMENTS

RESULT 1

I64155 hypothetical protein HI0638 - Haemophilus influenzae (strain Rd KW20)

C/Species: Haemophilus influenzae

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #ext_change 09-Jul-2004

C/Accession: I64155

R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, S.H.; Smith, T.F.; Luv, L.I.; Glodek, A.; Kelley, J.M.; Waidman, J.D.; Hendrickson, R.C.; Fritch, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995

A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A/Reference number: A64000; PMID:95350630; PMID:7542800

A/Accession: I64155

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-205 <TRIG>

A/Cross-references: UNIPROT:P44796; UNIPARC:UPI000013A674; GB:U32747; GB:I42023; NID:91

A/Note: best homolog was a hypothetical protein from Escherichia coli

Query Match 100.0%; Score 117; DB 2; Length 205;

Best Local Similarity 100.0%; Pred. No. 2.4e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 NYHDIYALAGVCSAKLVHQLA 25

NYHDIYALAGVCSAKLVHQLA 23

RESULT 2 AH0139 conserved hypothetical protein YP01637 [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #ext_change 09-Jul-2004

C/Accession: AH0139

Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

Ell, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,

Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AH0139

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-208 <KIR>

A/Cross-references: UNIPROT:Q8ZFC6; UNIPARC:UPI000000CD828; GB:AL590842; PIDN:CAC90459.1

A/Accession: YP01637

Query Match 81.2%; Score 95; DB 2; Length 208;

Best Local Similarity 78.3%; Pred. No. 9e-08;

Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 4 NYDITTLALGICGSARLVQOLA 26

RESULT 3

S19211

yycfC protein - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C:Accession: S19211; A64858

R:Green, S.M.; Drabble, W.T.

submitted to the EMBL Data Library, May 1991

A:Description: Molecular analysis of the purB-phoP region of Escherichia coli K12.

A:Reference number: S19210

A:Accession: S19211

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <GRS>

A:Cross-references: UNIPROT:P25746; UNIPARC:UPI000013A673; EMBL:X59307; NID:g42582; PIDN

A:Experimental source: strain K-12

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: A64858

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-213 <BLAT>

A:Cross-references: UNIPARC:UPI000013A673; GB:AE000213; GB:U00096; NID:g1787371; PIDN:AA

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: yycfC

A:Start codon: GTG

Query Match 81.2%; Score 95; DB 2; Length 213;

Best Local Similarity 78.3%; Pred. No. 9.2e-08;

Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 4 NYDITTLALGICGSARLVQOLA 26

RESULT 4

D90829

hypothetical protein ECs1604 [imported] - Escherichia coli (strain O157:H7, substrain R1

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: D90829

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

Gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: D90829

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <HAY>

A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI000000D0AD2; GB:BA000007; PIDN:BA035027.1;

A:Experimental source: strain O157:H7, substrain R1MD 0509952

C:Genetics:

A:Gene: ECs1604

Query Match 81.2%; Score 95; DB 2; Length 213;
Best Local Similarity 78.3%; Pred. No. 9.2e-08;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 4 NYDITTLALGICGSARLVQOLA 26

RESULT 5

B85687

hypothetical protein yycfC [imported] - Escherichia coli (strain O157:H7, substrain EDL93

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: B85687

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glanier, J.D.; Rose, D.J.; Mayhew

Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: B85687

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 <STO>

A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI000000D0AD2; GB:AE005174; NID:g12514786; P

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: yycfC

Query Match 81.2%; Score 95; DB 2; Length 213;
Best Local Similarity 78.3%; Pred. No. 9.2e-08;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 4 NYDITTLALGICGSARLVQOLA 26

RESULT 6

AB0647

conserved hypothetical protein STY1273 [imported] - Salmonella enterica subsp. enterica

C:Species: Salmonella enterica subsp. enterica serovar Typh

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AB0647

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrer,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AB0647

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <PAR>

A:Cross-references: UNIPARC:UPI0000059F77; GB:AL513382; PIDN:CAD08357.1; PID:g16502402;

C:Genetics:

A:Gene: STY1273

Query Match 78.6%; Score 92; DB 2; Length 215;
Best Local Similarity 73.9%; Pred. No. 2.8e-07;
Matches 17; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 6 NYDITTLALGICGSARLVQOLA 28

RESULT 7

D84960

hypothetical protein [imported] - Buchnera sp. (strain APS)

C:Species: Buchnera sp.

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001

C:Accession: D84960

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A

A:Reference number: A84930; MUID:20445173; PMID:10993077

A:Accession: D84960

A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-211 <STO>
 A/Cross-references: UNIPARC:UPI000005EAF; GB:AF000398; GSPDB:GN00144
 A/Experimental source: strain APS
 C/Genetics:
 A/Gene: ycfC; BU262

Query Match 61.5%; Score 72; DB 2; Length 211;
 Best Local Similarity 71.4%; Pred. No. 0.00049;
 Matches 15; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 HDIVLALAGVCGSAGLVHQLA 23
 DB 5 HLTITSLAGTCGSAHLVQLA 25

RESULT 8

B82237
 conserved hypothetical protein VC1127 [imported] - *Vibrio cholerae* (strain N16961 serogroup C/Species: *Vibrio cholerae*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C/Accession: B82237
 R/Residues: J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Swinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragol, I.; Sellers, F. 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A/Reference number: A82035; MWID:20406833; PMID:10952301

A/Accession: B82237

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-205 <HEI>

A/Cross-references: UNIPROT:Q9KX9; UNIPARC:UPI00000C2EB4; GB:AE004193; GB:AE003852; NIH

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

C/Genetics:

A/Gene: VC1127

A/Map position: 1

Query Match 47.9%; Score 56; DB 2; Length 205;
 Best Local Similarity 47.6%; Pred. No. 0.19;
 Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 HDIVLALAGVCGSAGLVHQLA 23
 DB 6 YDRTIAPAGTCQAVLVQVLA 26

RESULT 9

B82680
 conserved hypothetical protein XF1439 [imported] - *Xylella fastidiosa* (strain 9a5c)

C/Species: *Xylella fastidiosa*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C/Accession: B82680

R/anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A/Reference number: A82515; MWID:20365717; PMID:10910347

A/Note: for a complete list of authors see reference number A59328 below

A/Accession: B82680

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-204 <STM>

A/Cross-references: UNIPROT:Q9PDE0; UNIPARC:UPI00000C2723; GB:AE003974; GB:AE003849; NIH

A/Experimental source: strain 9a5c

R/Simpson, M.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B

Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincant, A.P.; Ferreira, A.J.S.

submitted to Genbank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; From

J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, U.P.; Krieger, J.E.; Kurame, E.E.; Laiz

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E

A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

, P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaak
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silvelli
 M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A/Reference number: A59328

A/Contents: annotation

C/Genetics:

A/Gene: XF1439

Query Match 41.0%; Score 48; DB 2; Length 204;
 Best Local Similarity 61.1%; Pred. No. 3.7;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 VIALAGVCGSAGLVHQLA 23
 DB 9 VIALAGVCGSAGLVHQLA 26

RESULT 10

C95338
 hypothetical protein Sma131 [imported] - *Sinorhizobium meliloti* (strain 1021) magaplae

C/Species: *Sinorhizobium meliloti*
 C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C/Accession: C95338

R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A/Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meli*

A/Reference number: A95262; MWID:21396509; PMID:1148132

A/Accession: C95338

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-531 <KUR>

A/Cross-references: UNIPROT:Q92281; UNIPARC:UPI00000CB152; GB:AE006469; PIDN:PAK65269.1

A/Experimental source: strain 1021, megaplasmaid pSymb

R/Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

peila, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federpiet, N.A.; Fisher, R.F.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Leleure

hebaul, P.; Vandenbol, M.; Vorholter, P.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K

A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.

A/Reference number: A96039; MWID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:

A/Gene: Sma131

A/Genome: plasmid

Query Match 40.6%; Score 47.5; DB 2; Length 531;
 Best Local Similarity 38.1%; Pred. No. 11;
 Matches 8; Conservative 9; Mismatches 3; Indels 1; Gaps 1;

QY 2 YHDIVLALAGVCGSAGLVHQL 22
 DB 334 FH-IVLALAGVCGSAGLVHQL 353

RESULT 11

AH0891
 adenyl-transferase [imported] - *Salmonella enterica* subsp. *enterica* serovar Typh

C/Species: *Salmonella enterica* subsp. *enterica* serovar Typh

A/Note: this species has also been called *Salmonella typhi*

C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C/Accession: AH0891

R/Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher

th, T.; Connerton, P.; Croft, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar

, S.; Mout, S.; O'Gea, P.

Nature 413, 848-852, 2001

A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A/Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* sero

A/Accession: AH0891

A/Status: preliminary

A:Molecule type: DNA
A:Residues: 1-947 <PAR>
A:Cross-references: UNIPARC:UPI000005A4A4; GB:AL513382; PIDN:CAD07726.1; PID:G16504278;
A:Genetics:
A:Gene: STY3380

Query Match 39.3%; Score 46; DB 2; Length 947;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 ALAGVCGSAGLVHQL 22
|||
DB 76 ALAGVCGDASLMREL 90

RESULT 12
T21782
hypochemical protein F35E2.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21782

R:Lennard, N.
submitted to the EMBL Data Library, November 1996

A:Reference number: 219471

A:Accession: T21782

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1265 <WIL>

A:Cross-references: UNIPROT:O62231; UNIPARC:UPI000007E856; EMBL:Z81528; PIDN:CAB04288.1;
A:Experimental source: clone F35E2

C:Genetics:

A:Gene: CESP:F35E2.9

A:Map position: 1

A:Introns: 41/1; 180/1; 250/1; 317/1; 379/2; 456/1; 520/1; 638/1; 709/2; 780/1; 850/1; 9

Query Match 39.3%; Score 46; DB 2; Length 1265;
Best Local Similarity 46.2%; Pred. No. 46;

Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 YHDIYALAGVCGQ 14
|:|:|:|:|:|:
DB 529 YNDVICTLAGICR 541

RESULT 13

T50513
hypochemical protein T27115_10 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004

C:Accession: T50513

R:Cholera, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Se

submitted to the Protein Sequence Database, May 2000

A:Reference number: 225102

A:Accession: T50513

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1857 <CHO>

A:Cross-references: UNIPROT:Q9LEX9; UNIPARC:UPI00000A3F58; EMBL:AL358732

A:Experimental source: cultivar Columbia; BAC clone T27115

C:Genetics:

A:Map position: 3

A:Introns: 295/1; 320/1; 550/3; 567/1; 721/3; 766/3; 801/3; 839/3; 865/2; 892/3; 1001/3;
A:Note: T27115_10

Query Match 39.3%; Score 46; DB 2; Length 1857;
Best Local Similarity 42.1%; Pred. No. 66;

Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 DIVIALAGVCGSAGLVHQL 22
::|:|:|:|:|:|:
DB 408 ELVAAIVSLCOSQKINHAL 426

RESULT 14

B86773
malate dehydrogenase (oxaloacetate-decarboxylating) (EC 1.1.1.38) mae [imported] - Lacto

N/Alternate names: malate oxidoreductase

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 05-Oct-2004

C:Accession: B86773

R:Bojorcin, A.; Wincker, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissbach, J.; Ehrli,

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: B86773

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <STO>

A:Cross-references: UNIPROT:Q9CGB2; UNIPARC:UPI00000C69D1; GB:AE005176; PID:G12724153; P

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: mae

C:Superfamily: malic enzyme

C:Keywords: oxidoreductase

Query Match 38.9%; Score 45.5; DB 2; Length 380;
Best Local Similarity 40.0%; Pred. No. 17;
Matches 10; Conservative 6; Mismatches 4; Indels 5; Gaps 1;

QY 2 YHD-----IVIALAGVCGSAGLVHQL 21
|||
DB 156 YHDDQAGTAVLVAGLMMAPLVNK 180

RESULT 15

T43916
chitinase A [imported] - Pyrococcus kodakaraensis

C:Species: Pyrococcus kodakaraensis

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C:Accession: T43916

R:Tanaka, T.; Fujiwara, S.; Nishikori, S.; Fukui, T.; Takagi, M.; Imanaka, T.

Appl. Environ. Microbiol. 65, 5338-5344, 1999

A:Title: A unique chitinase with dual active sites and triiple substrate binding sites fr

A:Reference number: 222722; MUID:20049967; PMID:10583986

A:Accession: T43916

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1215 <TMN>

A:Cross-references: UNIPROT:Q9UWR7; UNIPARC:UPI00000624AB; EMBL:AB024740; NID:G6580044;

C:Genetics:

A:Gene: ch1A

Query Match 38.9%; Score 45.5; DB 2; Length 1215;
Best Local Similarity 40.0%; Pred. No. 53;
Matches 10; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

QY 4 DIVIALAG-----VCGSAGLVHQLA 23
|:|:|:|:|:|:|:|:|:|:|:|:
DB 980 DVITAFGAVGVPVLCQQAKTREQLA 1004

Search completed: January 23, 2006, 09:32:47
Job time : 11.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 Seconds
(without alignments)
267,334 Million cell updates/sec

Title: US-10-501-838A-1
Perfect score: 117
Sequence: 1 NYHDIYVLAGVCGQSAKLVIHQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	100.0	205	1 Y638_HAEIN	P44796 haemophilus
2	117	100.0	205	2 O4QMS9_HAE18	Q4qms9 haemophilus
3	95	81.2	204	2 O65VV5_MANGM	O65vv5 manheimia
4	95	81.2	208	1 Y1637_TERPE	O8f763 yersinia pe
5	95	81.2	208	2 O66903_YERPS	O8x736 escherichia
6	95	81.2	213	1 YCFC_ECOL6	O8f1b7 escherichia
7	95	81.2	213	1 YCFC_ECOL6	P25746 escherichia
8	95	81.2	213	1 YCFC_ECOL6	O831f8 shigella fl
9	95	81.2	213	1 YCFC_ECOL6	O9c3f8 pasteurella
10	95	81.2	203	1 Y1850_PASMU	O8z7h0 salmonella
11	92	78.6	213	1 YCFC_SALT	O8z7p5 salmonella
12	92	78.6	213	1 YCFC_SALT	O8z7p5 salmonella
13	92	78.6	213	1 YCFC_SALT	O8z7p5 salmonella
14	88	75.2	212	2 O6D48_BRMCT	O6d48b erwinia car
15	88	75.2	212	2 O570C1_SALCH	O570c1 salmonella
16	78	66.7	208	1 Y2805_FHOL	O7n3b4 photorhabdu
17	72	61.5	211	1 Y262_BUCAI	P57350 buchiera ap
18	72	61.5	215	1 Y1650_HABDU	QV1v35 haemophilus
19	72	61.5	217	1 Y243_BUCBP	O8x8x5 buchiera ap
20	56	47.9	205	1 Y1127_VIBCH	O8x8x5 vibrio chol
21	56	47.9	205	1 Y1129_VIBPA	O8x8x5 vibrio para
22	56	47.9	205	1 Y1129_VIBPA	O8x8x5 vibrio para
23	56	47.9	205	1 Y1129_VIBPA	O8x8x5 vibrio para
24	55	47.0	205	1 Y1342_VIBBV	O8x8x5 vibrio vuln
25	55	47.0	205	1 Y1342_VIBBV	O8x8x5 vibrio vuln
26	52.5	44.9	367	2 O5U8T4_TOBAC	O5u8t4 nicotiana t
27	52.5	44.9	367	2 O5U8T4_TOBAC	O5u8t4 nicotiana t
28	50	42.7	577	2 O4UVH4_XANCP	O4uvh4 xanthomonas
29	49	41.9	205	2 O6B8K9_XANCP	O6b8k9 xanthomonas
30	49	41.9	207	2 O6C119_PROPR	O6c119 photobacter
31	49	41.9	245	2 O60CA8_METCA	O60ca8 methylotoc
					O7vxt7 bordetella

32	49	41.9	245	2 O7W6Y2_BORPA	O7w6y2 bordetella
33	49	41.9	245	2 O7W6Y2_BORPA	O7w6y2 bordetella
34	49	41.9	372	2 O7PSM6_ANOGA	O7psm6 anophelis g
35	49	41.9	390	2 O4S165_TETNG	O4s165 tetrarodon n
36	48	41.0	204	1 Y1439_XYLPA	O9pde0 xyella fas
37	48	41.0	204	1 Y661_XYLFT	O9pde0 xyella fas
38	48	41.0	394	2 O61GAT_CABRR	O61gat caenorhabd
39	48	41.0	446	2 O4TB11_TETNG	O4tb11 tetrarodon n
40	48	41.0	526	2 O6A0L6_DESPS	O6a0l6 deulfofale
41	48	41.0	925	2 O6Q122_RAT	O6q122 rattus norv
42	48	41.0	985	2 O7G6H1_ORYSA	O7g6h1 oryza sativ
43	48	41.0	986	2 O8S6V9_ORYSA	O8s6v9 oryza sativ
44	47.5	40.6	531	2 O92281_RHIME	O92281 rhicobium m
45	47	40.2	206	2 O7NGC7_GLOVI	O7ngc7 gloeobacter

ALIGNMENTS

RESULT 1
Y638_HAEIN STANDARD; PRT; 205 AA.
ID Y638_HAEIN
AC P44796;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein H10638.
GN OrderedLocustNames=H10638;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellales; Haemophilus.
NCBI_TaxID=727;
RX NCBI_TaxID=727;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=RD / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA McKenney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Ruhlmann J.L., Geoghagen N.S.M.,
RA Gehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd-5."
RL Science 269:496-512 (1995).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL: U32747; AAC2298.1; -; Genomic DNA.
DR PIR: I64155; I64155.
DR TIGR: H10638;
DR HAMAP: MF_00695; -; 1.
DR InterPro: IPR007451; DUF489.
DR Pfam: PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 23235 MW; 81F31B0186BF82DA CRC64;
Query Match 100.0%; Score 117; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 2.8e-10; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;
OY 1 NYHDIYVLAGVCGQSAKLVIHQLA 23
DB 3 NYHDIYVLAGVCGQSAKLVIHQLA 25

Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIVLALAGVQSARLVHOLA 23
DB 4 NYDITLALAGVQSARLVQOLA 26

RESULT 5

Q66903_YERPS PRELIMINARY; PRT; 208 AA.

DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN OrderedLocustNames=YP02411;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer F.W., Lemerding J., Scoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnepusch J., Marceau M., Medigue C.,
RA Simonet M., Chantal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia B.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL, BX936398; CAH21669.1; -; Genomic_DNA.
DR SMR, Q66903.2-207.
DR InterPro, IPR007451; DUF489.
DR Pfam, PF04356; DUF489; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE499723 CRC64;

Query Match 81.2%; Score 95; DB 2; Length 208;
Best Local Similarity 78.3%; Pred. No. 9.8e-07;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIVLALAGVQSARLVHOLA 23
DB 4 NYDITLALAGVQSARLVQOLA 26

RESULT 6

YCP_C0057 STANDARD; PRT; 213 AA.

DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocustNames=z1861, EC91604;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RC STRAIN=0157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glaesner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Pofel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta R.T., Potamovasis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).

RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).

RC STRAIN=0157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.,
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, AE005174; ANG5958.1; -; Genomic_DNA.
CC EMBL, BA000007; BAB35027.1; -; Genomic_DNA.
CC PIR, B85687; B85687.
CC PIR, D90829; D90829.
CC SMR, Q8X736; 2-213.
CC HMAP, MF_00695; 1.
CC InterPro, IPR007451; DUF489.
CC Pfam, PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22947 MW; E26F9678C3844E2 CRC64;

Query Match 81.2%; Score 95; DB 1; Length 213;
Best Local Similarity 78.3%; Pred. No. 1e-06;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIVLALAGVQSARLVHOLA 23
DB 4 NYDITLALAGVQSARLVQOLA 26

RESULT 7

YCP_EC06 STANDARD; PRT; 213 AA.

DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocustNames=c1511;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
RC STRAIN=06:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G., III, Redford P., Koesech P.,
RA Raeko D., Buckles B.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, AB016759; AAN79980.1; ALT_INIT; Genomic_DNA.
CC SMR, Q8FIB7; 2-213.

DR SMR; Q83LFB; 2-213.
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 213 AA; 22894 MW; 080BEC698C29FB4D CRC64;

Query Match 81.2%; Score 95; DB 1; Length 213;
 Best Local Similarity 78.3%; Pred. No. 1e-06;
 Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCSAKLVHQLA 23
 ||:|||||:|||||:|||||
 DB 4 NYHDIYALAGVCSAKLVHQLA 26

RESULT 10
 Y1850_PASMU STANDARD; PRT; 203 AA.

AC Q9CJY8;
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE Hypothetical UPF0274 protein PM1850.
 DR Hypothetical UPF0274 protein PM1850.
 GN OrderedLocustNames=PM1850;
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxId=747;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittem T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida PM70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; A806923; AA03934.1; -; Genomic_DNA.
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 203 AA; 22717 MW; 5EB91EBF0E181EDF CRC64;

Query Match 78.6%; Score 92; DB 1; Length 203;
 Best Local Similarity 78.3%; Pred. No. 2.9e-06;
 Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCSAKLVHQLA 23
 ||:|||||:|||||:|||||
 DB 3 NYHDIYALAGVCSAKLVHQLA 25

RESULT 11
 YCFC_SALTY STANDARD; PRT; 213 AA.

AC Q827H0;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein YCFC.
 GN Name=ycfc; OrderedLocustNames=STV1273, C1687;
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxId=601;

RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 RA Parthill J., Dougan G., James K.D., Thomson N.R., Pichard D., Main J.,
 RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin P., Haque A., Hien T.T., Holroyd S., Jagsis K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra F., Parry C.,
 RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).

RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=TY2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RX DOI=10.1128/JB.185.7.2330-2337.2003;
 RX Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).

CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; AL627269; CAD08357.1; ALT_INIT; Genomic_DNA.
 DR EMBL; AE016839; AA06312.1; ALT_INIT; Genomic_DNA.
 DR SMR; Q827H0; 2-213.
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 213 AA; 22942 MW; 63192D1C8066ED58 CRC64;

Query Match 78.6%; Score 92; DB 1; Length 213;
 Best Local Similarity 73.9%; Pred. No. 3.1e-06;
 Matches 17; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCSAKLVHQLA 23
 ||:|||||:|||||:|||||
 DB 4 NYHDIYALAGVCSAKLVHQLA 26

RESULT 12
 YCFC_SALTY STANDARD; PRT; 213 AA.

AC Q82PZ5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein YCFC.
 GN Name=ycfc; OrderedLocustNames=STW1233;
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxId=602;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=UT2 / SSGC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
 RA McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney B.,
 RA Ryan B., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;

RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2.";
RL Nature 413:852-856(2001).
CC -1 SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL, AB008754; AAL20162.1; ALT_INIT; Genomic_DNA.
DR SMR; 082P25; 2-213.
DR StvGene: SG272727; yfcC.
DR HAMAP; MF_006955; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match 78.6%; Score 92; DB 1; Length 213;
Best Local Similarity 73.9%; Pred. No. 3.1e-06;
Matches 17; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 NYHDIYALAGVCSAKLVHOLA 23
||:|||||:|||||:|||||
DB 4 NYIDITLALSGICSAKRLVQOLA 26

RESULT 13
OSPMJ3 SALPA PRELIMINARY; PRT; 213 AA.
ID O5PMJ3 SALPA PRELIMINARY;
AC O5PMJ3_

DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein yfcC.
GN Name=yfcC; OrderedLocustNames=SPAL617;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150;
RX PubMed=15531862; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.B., Clifton S.W., Latreille P.,
RA Porollik S., Sabo A., Meyer R., Bieri T., Ozerky P., McElellan M.,
RA Hartline C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du P., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Naeh W., Miner T., Mink P., Florea L.,
RA Delnauant K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spleth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of *Salmonella enterica* that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL, CP000025; AAV77544.1; -; Genomic_DNA.
DR SMR; O5PMJ3; 2-213.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match 78.6%; Score 92; DB 2; Length 213;
Best Local Similarity 73.9%; Pred. No. 3.1e-06;
Matches 17; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 NYHDIYALAGVCSAKLVHOLA 23
||:|||||:|||||:|||||
DB 4 NYIDITLALSGICSAKRLVQOLA 26

RESULT 14

O6D4E8 ERWCT
ID O6D4E8 ERWCT PRELIMINARY; PRT; 212 AA.
AC O6D4E8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=BCA2443;
OS *Erwinia carotovora* (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mangall K.,
RA Atlin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagsi K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmon G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen *Erwinia carotovora* subsp. atroseptica and characterization of virulence factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
DR EMBL; BX950851; CAG75345.1; -; Genomic_DNA.
DR SMR; O6D4E8; 2-207.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 212 AA; 22892 MW; A67813005F32B150 CRC64;

Query Match 75.2%; Score 86; DB 2; Length 212;
Best Local Similarity 73.9%; Pred. No. 1.3e-05;
Matches 17; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 NYHDIYALAGVCSAKLVHOLA 23
||:|||||:|||||:|||||
DB 4 NYIDITLALSGICSAKRLVQOLA 26

RESULT 15
OS7OCL SALCH
ID OS7OCL SALCH PRELIMINARY; PRT; 215 AA.
AC OS7OCL;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein yfcC.
GN Name=yfcC; OrderedLocustNames=SC1184;
OS *Salmonella cholerae-suis* (*Salmonella enterica*).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of *Salmonella enterica* serovar *Choleraesuis*, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698(2005).
DR EMBL; AE017220; AAK5090.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 215 AA; 23157 MW; C153B20629FAB8BD CRC64;

Query Match 75.2%; Score 86; DB 2; Length 215;
Best Local Similarity 72.7%; Pred. No. 1.4e-05;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 NYHDIYALAGVCSAKLVHOLA 22

Tue Jan 24 08:04:09 2006

us-10-501-838a-1.rup

Page 7

DB 6 NYDITLALSGICQSAVLVOL 27

Search completed: January 23, 2006, 09:31:05
Job time : 62.7 secs

This Page Blank (uoft)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838A-1

Sequence: 1 NYHDIYALAGVCSAKLVHQLA 23

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: Published Applications AA Main:
1: /cgm2_6/prodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgm2_6/prodata/1/pubppa/US08_PUBCOMB.pep.*
3: /cgm2_6/prodata/1/pubppa/US09_PUBCOMB.pep.*
4: /cgm2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
5: /cgm2_6/prodata/1/pubppa/US10B_PUBCOMB.pep.*
6: /cgm2_6/prodata/1/pubppa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	100.0	23	4	US-10-665-184-1
2	117	100.0	23	5	US-10-942-300-1
3	117	100.0	23	5	US-10-501-838A-1
4	117	100.0	205	4	US-10-665-184-59
5	117	100.0	205	5	US-10-942-300-59
6	117	100.0	205	5	US-10-501-838A-53
7	114	97.4	23	4	US-10-665-184-29
8	114	97.4	23	5	US-10-942-300-29
9	114	97.4	23	5	US-10-501-838A-29
10	114	97.4	29	4	US-10-665-184-33
11	114	97.4	29	5	US-10-942-300-33
12	114	97.4	30	5	US-10-501-838A-33
13	95	81.2	23	4	US-10-665-184-3
14	95	81.2	23	5	US-10-942-300-3
15	95	81.2	23	5	US-10-501-838A-3
16	95	81.2	24	4	US-10-665-184-25
17	95	81.2	24	5	US-10-942-300-25
18	95	81.2	24	5	US-10-501-838A-25
19	95	81.2	25	5	US-10-501-838A-19
20	95	81.2	25	5	US-10-501-838A-23
21	95	81.2	26	5	US-10-501-838A-21
22	95	81.2	29	5	US-10-501-838A-18
23	95	81.2	30	4	US-10-665-184-22
24	95	81.2	30	5	US-10-942-300-22
25	95	81.2	30	5	US-10-501-838A-20
26	95	81.2	30	5	US-10-501-838A-22
27	95	81.2	31	4	US-10-665-184-35

28	95	81.2	31	5	US-10-942-300-35	Sequence 35, Appl
29	95	81.2	31	5	US-10-501-838A-35	Sequence 35, Appl
30	95	81.2	213	4	US-10-665-184-61	Sequence 61, Appl
31	95	81.2	213	5	US-10-942-300-61	Sequence 61, Appl
32	95	81.2	213	5	US-10-501-838A-55	Sequence 55, Appl
33	95	81.2	23	4	US-10-665-184-2	Sequence 2, Appl
34	95	81.2	23	5	US-10-942-300-2	Sequence 2, Appl
35	95	81.2	23	5	US-10-501-838A-2	Sequence 2, Appl
36	95	81.2	23	4	US-10-665-184-60	Sequence 60, Appl
37	95	81.2	203	5	US-10-942-300-60	Sequence 60, Appl
38	95	81.2	203	5	US-10-501-838A-54	Sequence 54, Appl
39	95	81.2	23	4	US-10-665-184-26	Sequence 26, Appl
40	95	81.2	23	5	US-10-942-300-26	Sequence 26, Appl
41	95	81.2	23	5	US-10-501-838A-26	Sequence 26, Appl
42	95	81.2	30	4	US-10-665-184-36	Sequence 36, Appl
43	95	81.2	30	5	US-10-942-300-36	Sequence 36, Appl
44	95	81.2	30	5	US-10-501-838A-36	Sequence 36, Appl
45	95	81.2	22	4	US-10-665-184-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-10-665-184-1
Sequence 1, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 23
TYPE: PRT
ORGANISM: haemophilus influenzae
US-10-665-184-1
Query Match 100.0%; Score 117; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 NYHDIYALAGVCSAKLVHQLA 23
OY |||||
RESULT 2
US-10-942-300-1
Sequence 1, Application US/10942300
Publication No. US2005016103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
TITLE OF INVENTION: Biological Barrier
FILE REFERENCE: 24348-503
CURRENT APPLICATION NUMBER: US/10/942,300
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 1
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: haemophilus influenzae
US-10-942-300-1

Query Match 100.0%; Score 117; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSAKLVHQLA 23
|||
Db 1 NYHDIYALAGVCSAKLVHQLA 23

RESULT 3

US-10-501-838A-1
; Sequence 1, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-501-838A-1

Query Match 100.0%; Score 117; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSAKLVHQLA 23
|||
Db 1 NYHDIYALAGVCSAKLVHQLA 23

RESULT 4

US-10-665-184-59
; Sequence 59, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Haemophilus influenzae

Qy 1 NYHDIYALAGVCSAKLVHQLA 23
|||
Db 1 NYHDIYALAGVCSAKLVHQLA 23

US-10-665-184-59

Query Match 100.0%; Score 117; DB 4; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSAKLVHQLA 23
|||
Db 3 NYHDIYALAGVCSAKLVHQLA 25

RESULT 5

US-10-942-300-59
; Sequence 59, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-942-300-59

Query Match 100.0%; Score 117; DB 5; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSAKLVHQLA 23
|||
Db 3 NYHDIYALAGVCSAKLVHQLA 25

RESULT 6

US-10-501-838A-53
; Sequence 53, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 53
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-501-838A-53

Query Match 100.0%; Score 117; DB 5; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY      1 NYHDIVTALAGVCQSACL VHQLA 23  
        |||||  
Db      3 NYHDIVTALAGVCQSACL VHQLA 25
```

Db 1 NYHDIYVLAAGVCGSARLVHQLA 23

RESULT 11
US-10-942-300-33

Sequence 33, Application US/10942300
Publication No. US20050136103A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel

FILE REFERENCE: 24348-503

TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-503

CURRENT FILING DATE: 2004-09-16

PRIOR APPLICATION NUMBER: 10/665,184

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 10/664,989

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 60/503,615

PRIOR FILING DATE: 2003-09-17

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.2

SEQ ID NO 33

LENGTH: 29

TYPE: PRT

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Acylated Penetrating Peptide

NAME/KEY: MISC FEATURE

LOCATION: (29)..(29)

OTHER INFORMATION: wherein Xaa is lysine having a free amino group that is acylated

US-10-942-300-33

Query Match 97.4%; Score 114; DB 5; Length 29;
Best Local Similarity 95.7%; Pred. No. 1.4e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYVLAAGVCGSARLVHQLA 23

Db 1 NYHDIYVLAAGVCGSARLVHQLA 23

RESULT 12
US-10-501-838a-33

Sequence 33, Application US/10501838A

Publication No. US20050215478A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel A.

FILE REFERENCE: 24348-501 NATL

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-501 NATL

CURRENT FILING DATE: 2004-07-19

PRIOR APPLICATION NUMBER: PCT/IB03/00968

PRIOR FILING DATE: 2003-02-07

PRIOR APPLICATION NUMBER: US 60/355,396

PRIOR FILING DATE: 2002-02-07

NUMBER OF SEQ ID NOS: 72

SOFTWARE: PatentIn version 3.2

SEQ ID NO 33

LENGTH: 30

TYPE: PRT

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: Synthetic; penetrating peptide

NAME/KEY: MISC FEATURE

LOCATION: (27)..(27)

OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino

OTHER INFORMATION: groups of the lysine residue

NAME/KEY: MISC FEATURE

LOCATION: (30)..(30)

OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino

OTHER INFORMATION: groups of the lysine residue

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (30)..(30)

OTHER INFORMATION: wherein another molecule can be coupled to the penetrating

OTHER INFORMATION: peptide via the free amino groups of the lysine residue

US-10-501-838a-33

Query Match 97.4%; Score 114; DB 5; Length 30;
Best Local Similarity 95.7%; Pred. No. 1.5e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYVLAAGVCGSARLVHQLA 23

Db 1 NYHDIYVLAAGVCGSARLVHQLA 23

RESULT 13
US-10-665-184-3

Sequence 3, Application US/10665184

Publication No. US20040146549A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel

FILE REFERENCE: 24348-501CIP

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-501CIP

CURRENT FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: PCT/IB03/00968

PRIOR FILING DATE: 2003-02-07

PRIOR APPLICATION NUMBER: 60/355,396

PRIOR FILING DATE: 2002-02-07

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.2

SEQ ID NO 3

LENGTH: 23

TYPE: PRT

ORGANISM: Escherichia coli

US-10-665-184-3

Query Match 81.2%; Score 95; DB 4; Length 23;
Best Local Similarity 78.3%; Pred. No. 1.1e-07;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYVLAAGVCGSARLVHQLA 23

Db 1 NYHDIYVLAAGVCGSARLVHQLA 23

RESULT 14
US-10-942-300-3

Sequence 3, Application US/10942300

Publication No. US20050136103A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel

FILE REFERENCE: 24348-503

TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-503

CURRENT FILING DATE: 2004-09-16

PRIOR APPLICATION NUMBER: 10/665,184

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 10/664,989

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 60/503,615

; PRIOR FILING DATE: 2003-09-17
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 3
 ; LENGTH: 23
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-942-300-3

Query Match 81.2%; Score 95; DB 5; Length 23;
 Best Local Similarity 78.3%; Pred. No. 1.1e-07;
 Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVQSAKLVHOLA 23
 ||:||:||:||:||:||:||||
 DB 1 NYDYITLALAGICQSAKLVHOLA 23

RESULT 15
 US-10-501-838A-3
 ; Sequence 3, Application US/10501838A
 ; Publication No. US20050215478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ben-Sasson, Shmuel A.
 ; APPLICANT: Cohen, Sinai
 ; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
 ; TITLE OF INVENTION: Biological Barrier
 ; FILE REFERENCE: 24348-501 NATL
 ; CURRENT APPLICATION NUMBER: US/10/501,838A
 ; CURRENT FILING DATE: 2004-07-19
 ; PRIOR APPLICATION NUMBER: PCT/IB03/00968
 ; PRIOR FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: US 60/355,396
 ; PRIOR FILING DATE: 2002-02-07
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 3
 ; LENGTH: 23
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-501-838A-3

Query Match 81.2%; Score 95; DB 5; Length 23;
 Best Local Similarity 78.3%; Pred. No. 1.1e-07;
 Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVQSAKLVHOLA 23
 ||:||:||:||:||:||:||||
 DB 1 NYDYITLALAGICQSAKLVHOLA 23

Search completed: January 23, 2006, 11:52:31
 Job time : 60 secs

This Page Blank (Page 1)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-2

Sequence: 1 NYVDITLALAGVCAAKLVQCPA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: Genesegp1980a:***
2: genesegp1990a:***
3: genesegp2000a:***
4: genesegp2001a:***
5: genesegp2002a:***
6: genesegp2003a:***
7: genesegp2003b:***
8: genesegp2004a:***
9: genesegp2005a:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	23	ADBI6887	Adbi6887 Pasteurel
2	116	100.0	23	AEBO8219	Aebo8219 Pasteurel
3	116	100.0	203	AEBO8262	Aebo8262 Pasteurel
4	107	92.2	217	ABOC1272	Aboc1272 Klebsiell
5	103	88.8	23	ADBI6888	Adbi6888 Escherich
6	103	88.8	23	AEBO8220	Aebo8220 Escherich
7	103	88.8	24	ADBI6908	Adbi6908 Escherich
8	103	88.8	24	AEBO8242	Aebo8242 Escherich
9	103	88.8	25	ADBI6922	Adbi6922 Escherich
10	103	88.8	25	ADBI6906	Adbi6906 Escherich
11	103	88.8	26	ADBI6904	Adbi6904 E.coli pe
12	103	88.8	29	ADBI6921	Adbi6921 E.coli pe
13	103	88.8	30	ADBI6903	Adbi6903 E.coli pe
14	103	88.8	30	ADBI6905	Adbi6905 Escherich
15	103	88.8	30	AEBO8239	Aebo8239 Penetrati
16	103	88.8	31	ADBI6918	Adbi6918 Escherich
17	103	88.8	31	AEBO8252	Aebo8252 Penetrati
18	103	88.8	213	AEBO8263	Aebo8263 Escherich
19	97	83.6	23	ADBI6909	Adbi6909 Escherich
20	97	83.6	23	AEBO8243	Aebo8243 Escherich
21	97	83.6	30	ADBI6919	Adbi6919 Escherich
22	97	83.6	30	AEBO8253	Aebo8253 Penetrati
23	92	79.3	23	ADBI6886	Adbi6886 Haemophil
24	92	79.3	23	AEBO8218	Aebo8218 Haemophil

25	92	79.3	205	4	AAM50230
26	92	79.3	205	9	AEBO8261
27	90	77.6	22	6	ADBI6910
28	90	77.6	22	9	AEBO8244
29	90	77.6	29	6	ADBI6920
30	90	77.6	29	9	AEBO8254
31	89	76.7	23	6	ADBI6912
32	89	76.7	23	6	AEBO8246
33	89	76.7	29	6	ADBI6916
34	89	76.7	29	9	AEBO8250
35	86	74.1	209	6	ABM67069
36	85	73.3	216	7	ADFO7708
37	82	62.1	23	6	ADBI6889
38	72	62.1	23	9	AEBO8221
39	72	62.1	204	9	AEBO8264
40	71	61.2	23	6	ADBI6890
41	71	61.2	23	9	AEBO8222
42	71	61.2	211	9	AEBO8265
43	51	44.0	313	4	AAU08337
44	51	44.0	313	6	ABG73541
45	51	44.0	313	6	ABU62487

ALIGNMENTS

RESULT 1
ADBI6887
ID ADBI6887 standard; peptide; 23 AA.

XX ADBI6887;

XX 20-NOV-2003 (first entry)

XX Pasteurella multocida PM1850 penetrating peptide 2.

XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;

XX infertility; hormone; vitamin deficiency; neurodegenerative;

XX cardiovascular; haematological; endocrine disorder; obesity;

XX neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;

XX osteopathic; cytostatic; nootropic.

XX Pasteurella multocida.

XX WO2003066859-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003WO-IB000968.

XX 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Ben-Sasson SA, Cohen E;

XX WPI; 2003-697452/66.

XX New penetrating peptide, useful for preparing a composition for treating

XX or preventing e.g. endocrine disorders.

XX Claim 2; Page 14; 60pp; English.

XX This invention relates to a novel peptide sequences capable of

XX translocating across a biological barrier. Furthermore, it refers to

XX methods that use these peptides to facilitate penetration of a

XX biologically active effector molecule such as a drug or other therapeutic

XX agent across biological barriers e.g. epithelial or endothelial cells

XX sealed by tight junctions. This peptide is derived from a bacterial

XX toxin, an integral membrane or extracellular protein and can comprise an

XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin

XX or enzyme. The effector molecule, however, can comprise for example

XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from PM1850 of
 CC Pasteurella multocida and is penetrating peptide 2 of the invention.
 XX

SQ Sequence 23 AA;
 Query March 100.0%; Score 116; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.5e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDYITLALAGVCOAAKLVQGF 23
 1 NYDYITLALAGVCOAAKLVQGF 23

DB 1 NYDYITLALAGVCOAAKLVQGF 23

RESULT 2
 AEB08219
 ID AEB08219 standard; peptide; 23 AA.
 AC AEB08219;
 DT 25-AUG-2005 (first entry)
 DE Pasteurella multocida PM1850 penetrating peptide 2, SEQ ID NO: 2.
 XX

KM Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KM antidiabetic; endocrine disease; gastrointestinal disease;
 KM metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KM degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KM neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KM neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KM dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KM anticonvulsant; genetic disorder; cardiovascular disease;
 KM cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KM coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KM nutritional disorder; vitamin deficiency; renal disease; nephrologic;
 KM genitourinary disease; hematological disease; anti-anemic; anemia;
 KM autoimmune disease; immunosuppressive; immune deficiency;
 KM immunostimulant; infectious disease; antimicrobial; infection;
 KM erectile dysfunction; andrology; major depressive disorder;
 KM antidepressant; psychiatric disorder; pain; analgesic;
 KM bacterial infection; antibacterial; viral infection; virocidic;
 KM fungal infection; fungicide; parasitic infection; antiparasitic;
 KM renal failure; antifertility; antineumatic; cyostatic;
 KM antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX

OS Pasteurella multocida.
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX
 XX 17-SEP-2003; 2003US-00665184.
 XX
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX
 XX (COHEN/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX
 XX WPI, 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 46; SEQ ID NO 2; 59pp; English.
 XX

PS The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for transepithelial delivery of effector. The
 XX invention comprises the effector sequentially coupled with a counter ion
 XX and at least one hydrophobic agent, where the effector is selectively
 XX encapsulated into a complex. The invention is useful for translocating
 XX effectors across a biological barrier such as epithelial cells and
 XX endothelial cells during treatment and prevention of disease or
 XX pathological conditions (including endocrine disorders, diabetes,
 XX infertility, hormone deficiencies, osteoporosis, ophthalmological
 XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 XX Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 XX coagulable states, coronary disease, cerebrovascular events, metabolic
 XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 XX hematological disorders, anemia of different entities, immunologic and
 XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
 XX infectious diseases, viral infections, bacterial infections, fungal
 XX infections, parasitic infections, neoplastic diseases, multi-factorial
 XX disorders, impotence, chronic pain, depression, different fibrosis states
 XX and short stature) and for mucosal vaccination against anthrax and
 XX hepatitis B. The present sequence is the Pasteurella multocida PM1850
 XX penetrating peptide. This sequence is used in the effective translocation
 XX of aminoglycoside antibiotics and antifungal agents across an epithelial
 XX barrier.
 XX

SQ Sequence 23 AA;
 Query March 100.0%; Score 116; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.5e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDYITLALAGVCOAAKLVQGF 23
 1 NYDYITLALAGVCOAAKLVQGF 23

DB 1 NYDYITLALAGVCOAAKLVQGF 23

RESULT 3
 AEB08262
 ID AEB08262 standard; protein; 203 AA.
 AC AEB08262;
 DT 25-AUG-2005 (first entry)
 DE Pasteurella multocida protein, SEQ ID NO: 60.
 XX

KM Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KM antidiabetic; endocrine disease; gastrointestinal disease;
 KM metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KM degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KM neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KM neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KM dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KM anticonvulsant; genetic disorder; cardiovascular disease;
 KM cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KM coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KM nutritional disorder; vitamin deficiency; renal disease; nephrologic;
 KM genitourinary disease; hematological disease; anti-anemic; anemia;
 KM autoimmune disease; immunosuppressive; immune deficiency;
 KM immunostimulant; infectious disease; antimicrobial; infection;
 KM erectile dysfunction; andrology; major depressive disorder;
 KM antidepressant; psychiatric disorder; pain; analgesic;
 KM bacterial infection; antibacterial; viral infection; virocidic;
 KM fungal infection; fungicide; parasitic infection; antiparasitic;
 KM renal failure; antifertility; antineumatic; cyostatic;
 KM antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX

OS Pasteurella multocida.
 XX
 XX

PN US2005136103-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 16-SEP-2004; 2004US-00942300.
 XX
 PR 17-SEP-2003; 2003US-00664989.
 XX
 PR 17-SEP-2003; 2003US-00665184.
 XX
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 XX
 PI (COHEN/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 XX
 DR WPI, 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Disclosure; SEQ ID NO 60; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Pasteurella multocida protein
 CC containing penetrating peptide at N-terminal end.
 XX
 SQ Sequence 203 AA;
 XX
 Query Match 100.0%; Score 116; DB 9; Length 203;
 Best Local Similarity 100.0%; Pred. No. 1.8e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 NYDDITLALAGVCGAARLVQOFA 23
 DB 3 NYDDITLALAGVCGAARLVQOFA 25
 XX
 RESULT 4
 ID ABO61272 standard; protein; 217 AA.
 XX
 AC ABO61272;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polypeptide seq'd 7789.
 XX
 KM Recombinant expression vector; transcription regulatory element;
 XX Klebsiella pneumoniae protein; antibacterial; Vaccine.
 OS
 XX Klebsiella pneumoniae.
 XX
 PN US610836-B1.

XX
 PD 26-AUG-2003.
 XX
 PD 27-JAN-2000; 2000US-00489039.
 XX
 PF 29-JAN-1999; 99US-0117747P.
 XX
 PR (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PA Breton GL, Osborne M;
 XX
 PI WPI: 2003-895346/82.
 XX
 DR N-PSDB; ACH94823.
 XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 XX preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 PS Disclosure; SEQ ID NO 7789; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 SQ Sequence 217 AA;
 XX
 Query Match 92.2%; Score 107; DB 7; Length 217;
 Best Local Similarity 91.3%; Pred. No. 5.8e-09;
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 NYDDITLALAGVCGAARLVQOFA 23
 DB 8 NYDDITLALAGVCGAARLVQOLA 30
 XX
 RESULT 5
 ID ADB16888 standard; peptide; 23 AA.
 XX
 AC ADB16888;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Escherichia coli VCFC penetrating peptide 3.
 XX
 KM Penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 XX infertility; hormone; vitamin deficiency; neurodegenerative;
 KM cardiovascular; haematological; endocrine disorder; obesity;
 KM neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 XX osteopathic; cytoskeletal; nootropic.
 XX
 OS Escherichia coli.
 XX
 PN WO2003066859-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 07-FEB-2003; 2003WO-1B000968.
 XX
 PR 07-FEB-2002; 2002US-0355396P.
 XX
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 PI Ben-Sasson SA, Cohen E;
 XX
 DR WPI: 2003-697452/66.
 XX
 PT New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.

PS Claim 2; Page 14; 60pp; English.

CC This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dargirin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antihypertensive, osteopathic,
CC cyostatic or neurotropic activities. This peptide is from YCFC of
CC Escherichia coli and is penetrating peptide 3 of the invention.

XX Sequence 23 AA:

SO

Query Match 88.8%; Score 103; DB 6; Length 23;
Best Local Similarity 82.6%; Pred. No. 2e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDITTLAAGVCGAARLVQQA 23
DB 1 NYDITTLAAGVCGAARLVQQA 23

RESULT 6
AEB08220
ID AEB08220 standard; peptide; 23 AA.

XX
XX
XX AEB08220;
XX
XX 25-AUG-2005 (first entry)

DE Escherichia coli YCFC penetrating peptide 3, SEQ ID NO: 3.

XX
XX
XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
XX antidiabetic; endocrine disease; gastrointestinal disease;
XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;
XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
XX anticoagulant; genetic disorder; cardiovascular disease;
XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
XX genitourinary disease; hematological disease; antihaemic; anemia;
XX autoimmune disease; immunosuppressive; immune deficiency;
XX immunostimulant; infectious disease; antimicrobial; infection;
XX erectile dysfunction; andrology; major depressive disorder;
XX antidepressant; psychiatric disorder; pain; analgesic;
XX bacterial infection; antibacterial; viral infection; virocidic;
XX fungal infection; fungicide; parasitic infection; antiparasitic;
XX renal failure; antifertility; antipneumatic; cyostatic;
XX antinflammatory; hepatotropic; hepatitis B virus infection.

XX
XX Escherichia coli.
XX
XX US2005136103-A1.
XX
XX 23-JUN-2005.
XX
XX 16-SEP-2004; 2004US-00942300.
XX
XX 17-SEP-2003; 2003US-00664989.
XX
XX PR

PR 17-SEP-2003; 2003US-00665184.
PR 17-SEP-2003; 2003US-0503615P.
XX
XX (BENS/) BEN-SASSON S A.
XX (COHE/) COHEN E.
XX
XX Ben-Sasson SA, Cohen E;
XX WPI; 2005-444089/45.
XX
XX Composition used for translocating effectors across barrier such as
XX epithelial cells during treatment of e.g. endocrine disorders comprises
XX effector sequentially coupled with counter ion and hydrophobic agent.

PS Claim 53; SEQ ID NO 3; 59pp; English.

XX
XX
XX The present invention relates to a pharmaceutical composition of
XX penetrating peptides for transepithelial delivery of effector. The
XX invention comprises the effector sequentially coupled with a counter ion
XX and at least one hydrophobic agent, where the effector is selectively
XX encapsulated into a complex. The invention is useful for translocating
XX effectors across a biological barrier such as epithelial cells and
XX endothelial cells during treatment and prevention of disease or
XX pathological conditions (including endocrine disorders, diabetes,
XX infertility, hormone deficiencies, osteoporosis, ophthalmological
XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
XX Parkinson's disease, multiple sclerosis, Huntington's disease,
XX cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
XX coagulable states, coronary disease, cerebrovascular events, metabolic
XX disorders, obesity, vitamin deficiencies, renal disorder, renal failure,
XX hematological disorders, anemia of different entities, immunologic and
XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
XX infectious diseases, viral infections, bacterial infections, fungal
XX infections, parasitic infections, neoplastic diseases, multi-factorial
XX disorders, impotence, chronic pain, depression, different fibrosis states
XX and short stature) and for mucosal vaccination against anthrax and
XX hepatitis B. The present sequence is the Escherichia coli YCFC
XX penetrating peptide. This sequence is used in the effective translocation
XX of aminoglycoside antibiotics and antifungal agents across an epithelial
XX barrier.

SO Sequence 23 AA;

Query Match 88.8%; Score 103; DB 9; Length 23;
Best Local Similarity 82.6%; Pred. No. 2e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDITTLAAGVCGAARLVQQA 23
DB 1 NYDITTLAAGVCGAARLVQQA 23

RESULT 7
ADB16908
ID ADB16908 standard; peptide; 24 AA.

XX
XX
XX ADB16908;
XX
XX 20-NOV-2003 (first entry)

DE Escherichia coli YCFC penetrating peptide 30.

XX
XX
XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
XX infertility; hormone; vitamin deficiency; neurodegenerative;
XX cardiovascular; haematological; endocrine disorder; obesity;
XX neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
XX osteopathic; cyostatic; neurotropic.

XX
XX Escherichia coli.
XX
XX OS
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX PR

XX 07-FEB-2003; 2003WO-1B000968.
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 XX MPI; 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 XX Claim 2; Page 15; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cytoskeletal or neurotrophic activities. This peptide is from YCFC of
 CC Escherichia coli and is penetrating peptide 30 of the invention.
 XX
 XX Sequence 24 AA:
 SQ
 Query Match 88.8%; Score 103; DB 6; Length 24;
 Best Local Similarity 82.6%; Pred. No. 2.1e-09;
 Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCAQAKLVQQA 23
 ID AEB08242 standard; peptide; 24 AA.
 DB 2 NYDITLALAGVCAQAKLVQQA 24

RESULT 8
 AEB08242
 XX
 XX AEB08242;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 XX Escherichia coli YCFC penetrating peptide 30, SEQ ID NO: 25.
 XX
 KW Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neoplastic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticoagulant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; arteriosclerosis;
 KW coronary artery disease; cardiac; vasodilator; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;

KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antiferility; antineuritic; cytoskeletal;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX Escherichia coli.
 OS
 XX US200516103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX
 XX 17-SEP-2003; 2003US-0066184.
 XX
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PA Ben-Sasson SA, Cohen E;
 PI MPI; 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 46; SEQ ID NO 25; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hypercoagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, renal failure,
 CC disorders, obesity, vitamin deficiencies, renal disorders, metabolic
 CC hematological disorders, anemia of different entities, immunologic and
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 XX Sequence 24 AA:
 SQ
 Query Match 88.8%; Score 103; DB 9; Length 24;
 Best Local Similarity 82.6%; Pred. No. 2.1e-09;
 Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCAQAKLVQQA 23
 ID ADB16922 standard; peptide; 25 AA.
 DB 2 NYDITLALAGVCAQAKLVQQA 24

RESULT 9
 ADB16922
 XX
 XX ADB16922;
 XX
 DT 20-NOV-2003 (first entry)

```

XX DE Escherichia coli peptide 3 linked to recombinant human insulin.
XX KW recombinant human insulin; epithelial; endothelial; tight junction;
XX KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
XX KW cardiovascular; haematological; endocrine disorder; obesity;
XX KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
XX KW osteopathic; cyostatic; nootropic; penetrating peptide; diabetes.
XX OS Synthetic.
XX OS Escherichia coli.
XX FH Key Location/Qualifiers
XX FT Peptide 1..23
XX FT /note= "Penetrating peptide 3"
XX FT Modified-site 25
XX FT /note= "Recombinant human insulin peptide coupled to the
XX FT C-terminus"
XX PN MO2003066859-A2.
XX PD 14-AUG-2003.
XX XX 07-FEB-2003; 2003WO-IB000968.
XX PF 07-FEB-2002; 2002US-0355396P.
XX PR 07-FEB-2002; 2002US-0355396P.
XX PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX PI Ben-Sasson SA, Cohen E;
XX XX WPI, 2003-697452/66.
XX DR WPI, 2003-697452/66.
XX PT New penetrating peptide, useful for preparing a composition for treating
XX PT or preventing e.g. endocrine disorders.
XX PS Example 8; Page 42; 60pp; English.
XX XX This invention relates to a novel peptide sequences capable of
XX CC translocating across a biological barrier. Furthermore, it refers to
XX CC methods that use these peptides to facilitate penetration of a
XX CC biologically active effector molecule such as a drug or other therapeutic
XX CC agent across biological barriers e.g. epithelial or endothelial cells
XX CC sealed by tight junctions. This peptide is derived from a bacterial
XX CC toxin, an integral membrane or extracellular protein and can comprise an
XX CC anticonvulsant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX CC or enzyme. The effector molecule, however, can comprise for example
XX CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
XX CC factors. The penetrating peptide is useful for the treatment of various
XX CC conditions including diabetes, infertility, hormone and vitamin
XX CC deficiencies, neurodegenerative, cardiovascular, haematological and
XX CC endocrine disorders, as well as obesity and neoplastic disease.
XX CC Accordingly, the peptides of this invention can be used in compositions
XX CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
XX CC cyostatic or nootropic activities. This peptide sequence consists of the
XX CC Escherichia coli penetrating peptide 3 coupled to recombinant human
XX CC insulin in the absence of a detachable linker peptide as a cleavage site,
XX CC used to deliver insulin across mucosal epithelia as a treatment for
XX CC diabetes, in an exemplification of the invention.
XX SQ Sequence 25 AA;
XX
XX Query Match 88.8%; Score 103; DB 6; Length 25;
XX Best Local Similarity 82.6%; Pred. No. 2.2e-09;
XX Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 1 NYDITLALAGVCOAKLVQOFA 23
DB 1 NYDITLALAGVCOAKLVQOFA 23

```

```

ADBI6906
XX ID ADBI6906 standard; peptide: 25 AA.
XX AC ADBI6906;
XX XX 20-NOV-2003 (first entry)
XX DE Escherichia coli peptide 3 linked to linearised insulin receptor peptide.
XX KW linearised insulin receptor; epithelial; endothelial; tight junction;
XX KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
XX KW cardiovascular; haematological; endocrine disorder; obesity;
XX KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
XX KW osteopathic; cyostatic; nootropic; penetrating peptide; diabetes.
XX OS Synthetic.
XX OS Escherichia coli.
XX FH Key Location/Qualifiers
XX FT Peptide 1..23
XX FT /note= "Penetrating peptide 3"
XX FT Modified-site 25
XX FT /note= "Linearised insulin receptor peptide coupled to
XX FT the C-terminus"
XX PN MO2003066859-A2.
XX PD 14-AUG-2003.
XX XX 07-FEB-2003; 2003WO-IB000968.
XX PF 07-FEB-2002; 2002US-0355396P.
XX PR 07-FEB-2002; 2002US-0355396P.
XX PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX PI Ben-Sasson SA, Cohen E;
XX XX WPI, 2003-697452/66.
XX DR WPI, 2003-697452/66.
XX PT New penetrating peptide, useful for preparing a composition for treating
XX PT or preventing e.g. endocrine disorders.
XX PS Example 9; Page 43; 60pp; English.
XX XX This invention relates to a novel peptide sequences capable of
XX CC translocating across a biological barrier. Furthermore, it refers to
XX CC methods that use these peptides to facilitate penetration of a
XX CC biologically active effector molecule such as a drug or other therapeutic
XX CC agent across biological barriers e.g. epithelial or endothelial cells
XX CC sealed by tight junctions. This peptide is derived from a bacterial
XX CC toxin, an integral membrane or extracellular protein and can comprise an
XX CC anticonvulsant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX CC or enzyme. The effector molecule, however, can comprise for example
XX CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
XX CC factors. The penetrating peptide is useful for the treatment of various
XX CC conditions including diabetes, infertility, hormone and vitamin
XX CC deficiencies, neurodegenerative, cardiovascular, haematological and
XX CC endocrine disorders, as well as obesity and neoplastic disease.
XX CC Accordingly, the peptides of this invention can be used in compositions
XX CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
XX CC cyostatic or nootropic activities. This peptide sequence consists of the
XX CC Escherichia coli penetrating peptide 3 coupled to a linearised insulin
XX CC receptor, used to deliver insulin across mucosal epithelia as a treatment
XX CC for diabetes, in an exemplification of the invention.
XX SQ Sequence 25 AA;
XX
XX Query Match 88.8%; Score 103; DB 6; Length 25;
XX Best Local Similarity 82.6%; Pred. No. 2.2e-09;
XX Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
OY 1 NYDITLALAGVCOAKLVQOFA 23

```

Db 1 NYDITLALAGICGASRLVQOLA 23

RESULT 11

ID ADB16904 standard; peptide; 26 AA.

AC ADB16904;

DT 20-NOV-2003 (first entry)

DE E_coli penetrating peptide 3 linked to heparin without a cleavage site.

KW heparin; epithelial; endothelial; tight junction; diabetes; infertility;
KW hormone; vitamin deficiency; neurodegenerative; cardiovascular;
KW haematological; endocrine disorder; obesity; neoplastic disease;
KW neuroprotective; cardiac; antiarteriosclerotic; osteopathic; cytostatic;
KW neurotropic; penetrating peptide.

OS Synthetic.
OS Escherichia coli.

FH Key Location/Qualifiers

FT Peptide 1..23

FT Modified-site 26 /note= "Penetrating peptide 3"

FT /note= "Heparin peptide coupled to the C-terminus"

PN WO2003066859-A2.

PD 14-AUG-2003.

PP 07-FEB-2003; 2003WO-1B000968.

PR 07-FEB-2002; 2002US-0355396P.

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PI Ben-Sasson SA, Cohen E;

DR WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating

PS or preventing e.g. endocrine disorders.

XX Example 10; Page 44; 60pp; English.

CC This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
CC cytostatic or neurotropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to heparin in the absence
CC of a detachable linker peptide as a cleavage site, used to deliver low
CC molecular weight heparin across mucosal epithelia in an exemplification
CC of the invention.

XX Sequence 26 AA;

Query Match 88.8%; Score 103; DB 6; Length 26;
Best Local Similarity 82.6%; Pred. No. 2.3e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 NYDITLALAGICGASRLVQOLA 23
DB 1 NYDITLALAGICGASRLVQOLA 23

RESULT 12

ID ADB16921 standard; peptide; 29 AA.

AC ADB16921;

DT 20-NOV-2003 (first entry)

DE E_coli peptide 3 linked to recombinant human insulin via cleavage site.

KW recombinant human insulin; epithelial; endothelial; tight junction;
KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
KW cardiovascular; haematological; endocrine disorder; obesity;
KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
KW osteopathic; cytostatic; neurotropic; detachable linker;
KW penetrating peptide; diabetes.

OS Synthetic.
OS Escherichia coli.

FH Key Location/Qualifiers

FT Peptide 1..23

FT Cleavage-site 26..29 /note= "Penetrating peptide 3"

FT Modified-site 29 /note= "Detachable linker peptide"

FT /note= "Recombinant human insulin peptide coupled to the
C-terminus"

PN WO2003066859-A2.

PD 14-AUG-2003.

PP 07-FEB-2003; 2003WO-1B000968.

PR 07-FEB-2002; 2002US-0355396P.

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PI Ben-Sasson SA, Cohen E;

DR WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating

PS or preventing e.g. endocrine disorders.

XX Example 8; Page 42; 60pp; English.

CC This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions

CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopachic,
CC cyrostatic or nootropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to recombinant human
CC insulin via the detachable linker peptide as a cleavage site, used to
CC deliver insulin across mucosal epithelia as a treatment for diabetes, in
CC an exemplification of the invention.
XX
SQ Sequence 29 AA;
Query Match 88.8%; Score 103; DB 6; Length 29;
Best Local Similarity 82.6%; Pred. No. 2.6e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 NYDITLALAGVCOAKLVQOFA 23
DB 1 NYDITLALAGVCOAKLVQOFA 23
RESULT 13
ADBI6903
ID ADBI6903 standard; peptide; 30 AA.
AC ADBI6903;
XX
DT 20-NOV-2003 (first entry)
XX
DE E_coli penetrating peptide 3 linked to heparin through a cleavage site.
XX
KW heparin; epithelial; endothelial; tight junction; diabetes; infertility;
KW hormone; vitamin deficiency; neurodegenerative; cardiovascular;
KW haematological; endocrine disorder; obesity; neoplastic disease;
KW neuroprotective; cardiant; antiarteriosclerotic; osteopachic; cyrostatic;
KW nootropic; detachable linker; penetrating peptide.
XX
OS Synthetic.
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /note= "Penetrating peptide 3"
FT Cleavage-site 26..29
FT /note= "Detachable linker peptide"
FT Modified-site 30
FT /note= "Heparin peptide coupled to the C-terminus"
XX
PN WO2003066859-A2.
XX
PD 14-AUG-2003.
XX
PP 07-FEB-2003; 2003WO-IB000968.
XX
PR 07-FEB-2002; 2002US-0355396P.
XX
PI (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Ben-Saason SA, Cohen E;
XX
XX WPI, 2003-697452/66.
XX
PT New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
PS Example 10; Page 44; 60pp; English.
XX
CC This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example

CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, delargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopachic,
CC cyrostatic or nootropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to heparin via the
CC detachable linker peptide as a cleavage site, used to deliver low
CC molecular weight heparin across mucosal epithelia in an exemplification
CC of the invention.
XX
SQ Sequence 30 AA;
Query Match 88.8%; Score 103; DB 6; Length 30;
Best Local Similarity 82.6%; Pred. No. 2.7e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 NYDITLALAGVCOAKLVQOFA 23
DB 1 NYDITLALAGVCOAKLVQOFA 23
RESULT 14
ADBI6905
ID ADBI6905 standard; peptide; 30 AA.
AC ADBI6905;
XX
DT 20-NOV-2003 (first entry)
XX
DE Escherichia coli peptide 3 coupled to imaging compound linker, IBM-002.
XX
KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
KW cardiant; antiarteriosclerotic; osteopachic; cyrostatic; nootropic;
KW imaging linker; penetrating peptide; IBM-002.
XX
OS Synthetic.
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 30
FT /note= "C-terminal amide"
XX
PN WO2003066859-A2.
XX
PD 14-AUG-2003.
XX
PP 07-FEB-2003; 2003WO-IB000968.
XX
PR 07-FEB-2002; 2002US-0355396P.
XX
PI (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Ben-Saason SA, Cohen E;
XX
XX WPI, 2003-697452/66.
XX
PT New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
PS Example 1; Page 33; 60pp; English.
XX
CC This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic

CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide sequence is IBM-002,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 XX
 SQ Sequence 30 AA;
 Query Match 88.8%; Score 103; DB 6; Length 30;
 Best Local Similarity 82.6%; Pred. No. 2.7e-09;
 Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYDITLALAGVCAAKLVQQA 23
 DB 1 NYDITLALAGVCAAKLVQQA 23
 RESULT 15
 AEB08239 standard; peptide; 30 AA.
 XX
 AC AEB08239;
 DT 25-ANG-2005 (first entry)
 DE Penetrating peptide SEQ: 22 used in composition for mucosal vaccination.
 XX
 KW Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; anemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Unidentified.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 1
 FT /note= "N-terminal acylated"
 FT Misc-difference 30
 FT /note= "Optionally C-terminal amide, optionally the free
 FT amino group of lysine is acylated with a fatty acid"
 XX
 PN US2005136103-A1.
 PD 23-JUN-2005.
 XX 16-SEP-2004; 2004US-00942300.

XX
 PR 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 DR MPI; 2005-44089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 PS Claim 63; SEQ ID NO 22; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Huntington's disease,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, cerebrovascular events, metabolic
 CC coagulable states, coronary disease, renal disorders, renal failure,
 CC disorders, obesity, vitamin deficiencies, anemia of different etiology, immune
 CC hematological disorders, autoimmune diseases, immune deficiencies,
 CC rheumatologic disorders, viral infections, bacterial infections, fungal
 CC infectious diseases, parasitic infections, neoplastic diseases, multi-factorial
 CC infections, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is a penetrating peptide (IBM-002) used
 CC in the composition for mucosal vaccination using a counter ion and a
 CC penetrating peptide.
 XX
 SQ Sequence 30 AA;
 Query Match 88.8%; Score 103; DB 9; Length 30;
 Best Local Similarity 82.6%; Pred. No. 2.7e-09;
 Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYDITLALAGVCAAKLVQQA 23
 DB 1 NYDITLALAGVCAAKLVQQA 23
 Search completed: January 23, 2006, 09:20:46
 Job time : 69.2 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 Seconds

(without alignments)
267.334 Million cell updates/sec

Title: US-10-501-838a-2

Perfect score: 116
Sequence: 1 NYDITLALAGVCOAKLVQDPA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: 1: uniprot_05.80:*

2: uniprot_aprot:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	203	1 Y1850_PASMU	09c3j8 pasteurrella
2	113	97.4	204	2 Q65VVS_MANSM	Q65VVS manheimia
3	103	88.8	208	1 Y1637_YERPE	08zf66 yersteinia pe
4	103	88.8	208	2 Q669Q3_YERPS	0669q3 yersteinia ps
5	103	88.8	213	1 YCFC_ECO57	08x736 escherichia
6	103	88.8	213	1 YCFC_ECO57	08f1d7 escherichia
7	103	88.8	213	1 YCFC_ECO57	08f1d7 escherichia
8	103	88.8	213	1 YCFC_ECO57	08f1d7 escherichia
9	100	86.2	213	1 YCFC_ECO57	08f1d7 escherichia
10	100	86.2	213	1 YCFC_ECO57	08f1d7 escherichia
11	100	86.2	213	1 YCFC_ECO57	08f1d7 escherichia
12	96	82.8	215	2 Q64B8_ERWCH	Q64b8 erwina car
13	96	82.8	215	2 Q64B8_ERWCH	Q64b8 erwina car
14	92	79.3	205	1 Y638_HABIN	P44796 haemophilus
15	92	79.3	205	2 Q4QMS9_HAB18	Q4qms9 haemophilus
16	90	77.6	215	1 Y1650_HABDU	Q7v135 haemophilus
17	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
18	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
19	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
20	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
21	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
22	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
23	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
24	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
25	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
26	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
27	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
28	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
29	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
30	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus
31	86	74.1	208	1 Y2805_PHOUL	Q7n3b4 photorhabdus

ALIGNMENTS

32	50	43.1	390	1	NK2R_RAT	P16610 rattus norv
33	49	42.2	130	2	Q96KEO_HUMAN	Q96keo homo sapien
34	49	42.2	237	2	Q96NDE_RHIZO	Q96nde rhizobium 1
35	49	42.2	276	2	Q9USG6_CHLIP	Q9usg6 chlamydia p
36	49	42.2	322	2	Q928Q5_CHLIP	Q928q5 chlamydia p
37	49	42.2	372	2	Q7PSM6_ANOGA	Q7psm6 anopheles g
38	49	42.2	384	1	NK2R_BOVIN	P05363 bos taurus
39	49	42.2	384	1	NK2R_CANFA	Q5dub2 canis fam1
40	49	42.2	398	1	NK2R_HUMAN	P21452 homo sapien
41	49	42.2	398	2	Q8NGQ8_HUMAN	Q8ngq8 homo sapien
42	49	42.2	398	2	Q9UDDE_HUMAN	Q9udee homo sapien
43	49	42.2	398	2	Q4QRI5_HUMAN	Q4qri5 homo sapien
44	49	42.2	402	1	NK2R_CAVPO	Q64077 cavia porce
45	49	42.2	599	2	Q726U4_DESVH	Q726u4 desulfovibr

RESULT 1

Y1850_PASMU STANDARD, PRT, 203 AA.

AC 09C3J8;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Hypothetical protein PF0274, protein P1850.

GN OrderedCusNames=PM1850;

OS Pasteurella multocida.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Pasteurella.

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

OC NCB1_TaxID=7477;

```
CC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
CC Pasteurellaceae; Mannheimia.
CX NCBI_TaxID=221988;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15378067; DOI=10.1038/nbt1010;
RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rhn J.-K.,
RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;
RT "The genome sequence of the capnophilic rumen bacterium Mannheimia
succiniciproducens."
RL Nat. Biotechnol. 22:1275-1281(2004).
DR EMBL; AE016827; AAU36905.1; -; Genomic_DNA.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 204 AA; 23036 MW; 3D64A70A30C33CCB CRC64;

Query Match 97.4%; Score 113; DB 2; Length 204;
Best Local Similarity 95.7%; Pred. No. 1.3e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDDITLALAGVCOAKLVQOFA 23
Db 4 NYDDITLALAGVCOAKLVQOFA 26

RESULT 3
Y1637_YERPE STANDARD; PRT; 208 AA.
ID Y1637_YERPE
AC Q82F06; Q8D001;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein YP01637/Y1798/YP167.
GN OrderedLocustNames=YP01637, Y1798, YP167;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
CX NCBI_TaxID=632;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Tiltball R.W., Holden M.T.G.,
RA Prentice M.B., Sebatia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN (2)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=1214230;
RX DOI=10.1126/SCIENCE.1184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.U., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.B., Brubaker R.R., Plano G.V.,
RA Strelley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN (3)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin H., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
```

```
RT avirulent to humans."
RL DNA Res. 11:179-197(2004).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AJ414149; CAC90459.1; -; Genomic_DNA.
DR EMBL; AE013782; AAM85366.1; ALT INIT; Genomic_DNA.
DR EMBL; AE017133; AAS61994.1; -; Genomic_DNA.
DR PIR; AH0199; AH0199.
DR SMR; Q82F06; 2-207.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
FT CONFLICT 1 M -> MIV (in Ref. 2).
SQ SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE49723 CRC64;

Query Match 88.8%; Score 103; DB 1; Length 208;
Best Local Similarity 82.6%; Pred. No. 5.3e-08;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDDITLALAGVCOAKLVQOFA 23
Db 4 NYDDITLALAGVCOAKLVQOFA 26

RESULT 4
Q66903_YERPS PRELIMINARY; PRT; 208 AA.
ID Q66903_YERPS
AC Q66903;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=YPTB2431;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
CX NCBI_TaxID=633;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel B., Larimer F.W., Jamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Himebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derblie A., Hauser L.U., Garcia B.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH21669.1; -; Genomic_DNA.
DR SMR; Q66903; 2-207.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE49723 CRC64;

Query Match 88.8%; Score 103; DB 2; Length 208;
Best Local Similarity 82.6%; Pred. No. 5.3e-08;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDDITLALAGVCOAKLVQOFA 23
Db 4 NYDDITLALAGVCOAKLVQOFA 26

RESULT 5
```

```

YCFC_ECO57
ID YCFC_ECO57 STANDARD; PRT; 213 AA.
AC 08X736;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocNames=21861, EC61604;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=83334;
(1)
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=0157:H7 / EDL933 / ATCC 700927 / EHEC;
RA MEDLINE=21074955; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.,
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
(2)
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=0157:H7 / Sakai / RIMD 0509952 / EHEC;
RA MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
-1- SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AB005174; AAC55958.1; -; Genomic DNA.
CC EMBL: BA000007; BAB35027.1; -; Genomic DNA.
CC PIR: B85687; B85687.
CC PIR: D90829; D90829.
CC SMR: O8X736; 2-213.
CC HAMAP: MF_00695; -; 1.
CC InterPro: IPR007451; DUF489.
CC Pfam: PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC KW SEQUENCE 213 AA; 22947 MW; E26EF9678C3844E2 CRC64;
SQ
Query Match 88.8%; Score 103; DB 1; Length 213;
Best Local Similarity 82.6%; Pred. No. 5,4e-08;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 NYSDITLALAGVCAAKLVQOFA 23
DB 4 NYSDITLALAGVCAAKLVQOFA 26

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=217992;
(1)
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=06.H1 / CFT073 / ATCC 700928 / UPEC;
RA MEDLINE=22288234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raebio D., Buckles E.L., Itoh S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AE016759; AAN79980.1; ALT_INIT; Genomic DNA.
CC SMR: Q8R1B7; 2-213.
CC HAMAP: MF_00695; -; 1.
CC InterPro: IPR007451; DUF489.
CC Pfam: PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC KW SEQUENCE 213 AA; 22934 MW; 71C8F96F8EAA37B8 CRC64;
SQ
Query Match 88.8%; Score 103; DB 1; Length 213;
Best Local Similarity 82.6%; Pred. No. 5,4e-08;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 NYSDITLALAGVCAAKLVQOFA 23
DB 4 NYSDITLALAGVCAAKLVQOFA 26

```

```
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to
CC framehifts.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: X59307; CAA41995.1; -; Genomic DNA.
CC EMBL: M74924; -; NOT ANNOTATED; CDS; Genomic DNA.
CC EMBL: U00096; AAC74216.1; -; Genomic DNA.
CC EMBL: D90748; BAA35954.1; -; Genomic DNA.
CC EMBL: D90749; BAA35963.1; -; Genomic DNA.
CC PIR: S19211; S19211.
CC PDB: 1QZ4; X-ray; A=2-213.
CC PDB: 1SDI; X-ray; A=2-213.
CC Ecobase; EBI321; -;
CC Ecobase; EBI345; ycfC.
CC HAMAP: MF_00695; -; 1.
CC InterPro: IPR007451; DUF489.
CC Pfam: PF04356; DUF489; 1.
CC 3D-structure; Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22948 MW; E26F9698C3CEB42 CRC64;

Query Match 88.8%; Score 103; DB 1; Length 213;
Best Local Similarity 82.6%; Pred. No. 5.4e-08;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAKLVQCPA 23
DB 4 NYDITLALAGICGASRLVQQLA 26

RESULT 8
YCF_C SHIFL STANDARD; PRT; 213 AA.
ID YCF_C SHIFL STANDARD; PRT; 213 AA.
AC O831F8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocustNames=SF1151, S1234;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
CX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=1384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kang B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
```

```
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AB005674; AAN42768.1; ALT INIT; Genomic DNA.
CC EMBL: AE016982; AAP16557.1; -; Genomic DNA.
CC SNR: O831F8; 2-213.
CC HAMAP: MF_00695; -; 1.
CC InterPro: IPR007451; DUF489.
CC Pfam: PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22894 MW; 0B0BEC698C29FB4D CRC64;

Query Match 88.8%; Score 103; DB 1; Length 213;
Best Local Similarity 82.6%; Pred. No. 5.4e-08;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAKLVQCPA 23
DB 4 NYDITLALAGICGASRLVQQLA 26

RESULT 9
YCF_C SALTI STANDARD; PRT; 213 AA.
ID YCF_C SALTI STANDARD; PRT; 213 AA.
AC O827H0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocustNames=STY1273, t1687;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
CX NCBI_TaxID=601;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1039/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltham T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jaseil K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
```

RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyanni V., Schwartz D.C., Blattner P.R.;
 RA "Comparative genomics of *Salmonella enterica* serovar Typhimurium Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, AL627269, CAD08357.1; ALT_INT; Genomic DNA.
 DR EMBL, AE016839, AAO69312.1; ALT_INT; Genomic DNA.
 DR SMR, Q827H0, 2-213.
 DR HAMAP, MF_00695, -; 1.
 DR InterPro, IPR007451, DUF489.
 DR Pfam, PF04356, DUF489, 1.
 KW Complete proteome; Hypothetical protein.
 SO SEQUENCE 213 AA; 22942 MW; 63192D1C806ED58 CRC64;
 NCBI_Taxid=602;
 Query Match 86.2%; Score 100; DB 1; Length 213;
 Best Local Similarity 78.3%; Pred. No. 1.6e-07;
 Matches 18; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYDITLALAGVCAKLVQQA 23
 DB 4 NYDITLALSGICQARLVQQLA 26
 ID YCRC_SALTY STANDARD; PRT; 213 AA.
 AC Q82E25;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein yCrc.
 GN Name=yCrc; OrderedLocNames=STM1233;
 OS *Salmonella typhimurium*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_Taxid=602;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=L72 / SGGC1412 / ATCC 700720.
 RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT L72.";
 RL Nature 413:852-856(2001).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, AB008754, AL20162.1; ALT_INT; Genomic DNA.
 DR SMR, Q82E25, 2-213.
 DR EYGene, SGT7777, yCrc.
 DR HAMAP, MF_00695, -; 1.
 DR InterPro, IPR007451, DUF489.
 DR Pfam, PF04356, DUF489, 1.
 KW Complete proteome; Hypothetical protein.
 SO SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match 86.2%; Score 100; DB 1; Length 213;
 Best Local Similarity 78.3%; Pred. No. 1.6e-07;
 Matches 18; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYDITLALAGVCAKLVQQA 23
 DB 4 NYDITLALSGICQARLVQQLA 26
 ID Q6DAE8_ERWCT PRELIMINARY; PRT; 212 AA.
 AC Q6DAE8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein
 GN OrderedLocNames=SCA443;
 OS *Erwinia carotovora* (subsp. atroseptica) (Pectobacterium atrosepticum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 OX NCBI_Taxid=29471;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SCRI 1043 / ATCC BAA-672;
 RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
 RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
 RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
 RA Akin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
 RA Frazer A., Hance Z., Hauser H., Jagers K., Moule S., Norbertczak H.,
 RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
 RT restricted serovars of *Salmonella enterica* that cause typhoid.";
 RL Nat. Genet. 36:1268-1274(2004).
 DR EMBL, CP000026, AAV77544.1; -; Genomic DNA.
 DR SMR, Q5PMJ3, 2-213.
 DR InterPro, IPR007451, DUF489.
 DR Pfam, PF04356, DUF489, 1.
 KW Complete proteome; Hypothetical protein.
 SO SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;
 Query Match 86.2%; Score 100; DB 2; Length 213;
 Best Local Similarity 78.3%; Pred. No. 1.6e-07;
 Matches 18; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYDITLALAGVCAKLVQQA 23
 DB 4 NYDITLALSGICQARLVQQLA 26
 ID Q6DAE8_ERWCT PRELIMINARY; PRT; 212 AA.
 AC Q6DAE8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein
 GN OrderedLocNames=SCA443;
 OS *Erwinia carotovora* (subsp. atroseptica) (Pectobacterium atrosepticum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 OX NCBI_Taxid=29471;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SCRI 1043 / ATCC BAA-672;
 RX PubMed=15263089; DOI=10.1038/ng1470;
 RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
 RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozerky P., Mclellan M.,
 RA Harting S., Wang C., Nguyen C., Berghoff A., Elliott G.,
 RA Kohlberg S., Strong C., Du F., Carter J., Kremliki C., Layman D.,
 RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Mink P.,
 RA Delehaunt J., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
 RA Spieth J., Wilson R.K.;
 RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
 RT restricted serovars of *Salmonella enterica* that cause typhoid.";
 RL Nat. Genet. 36:1268-1274(2004).
 DR EMBL, CP000026, AAV77544.1; -; Genomic DNA.
 DR SMR, Q5PMJ3, 2-213.
 DR InterPro, IPR007451, DUF489.
 DR Pfam, PF04356, DUF489, 1.
 KW Complete proteome; Hypothetical protein.
 SO SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

```

RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.,
RA "Genome sequence of the enterobacterial phytopathogen Erwinia
RA carotovora subsp. atroseptica and characterization of virulence
RA factor."
RA Proc. Natl. Acad. Sci. U.S.A. 101.11105-11110 (2004).
DR EMBL: BX950851; CAG75345.1; -; Genomic_DNA.
DR SMR: OGD488; 2-207.
DR InterPro: IPR007451; DUF489.
DR Pfam: PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 212 AA; 22892 MW; A67813005F32B150 CRC64;

Query Match
Best Local Similarity 82.8%; Score 96; DB 2; Length 212;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYVDITLALAGVCOAKLVQCPA 23
Db 4 NYVDITLALAGVCOAKLVQCPA 26

RESULT 13
ID Q57QC1 SALCH PRELIMINARY; PRT; 215 AA.
AC Q57QC1
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Hypothetical protein yjcC.
GN Name=yjcC; OrderedLocustNames=SC1184;
OS Salmonella cholerae-suis (Salmonella enterica).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RA "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RA highly invasive and resistant zoonotic pathogen."
RL Nucleic Acids Res. 33:1690-1698 (2005).
DR EMBL: AB017220; AAX65090.1; -; Genomic_DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 215 AA; 23157 MW; C153B20629FAE8BD CRC64;

Query Match
Best Local Similarity 82.8%; Score 96; DB 2; Length 215;
Matches 17; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYVDITLALAGVCOAKLVQCPA 21
Db 6 NYVDITLALAGVCOAKLVQCPA 26

RESULT 14
ID Y638_HAEIN STANDARD; PRT; 205 AA.
AC P44796;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein H10638.
GN OrderedLocustNames=H10638;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Rd / KW20 / ATCC 51907;

```

```

RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,
RA McInerney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Ureback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -1 SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: U32747; AAC22298.1; -; Genomic_DNA.
CC PIR: I64155; I64155.
CC DR TIGR: H10638; -.
CC DR HAMAP: MF_00695; -; 1.
CC DR InterPro: IPR007451; DUF489.
CC Pfam: PF04355; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 23235 MW; 81F31B0186BF82DA CRC64;

Query Match
Best Local Similarity 79.3%; Score 92; DB 1; Length 205;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NYVDITLALAGVCOAKLVQCPA 23
Db 3 NYVDITLALAGVCOAKLVQCPA 25

RESULT 15
ID Q4QMS9_HAE18 PRELIMINARY; PRT; 205 AA.
AC Q4QMS9;
DT 13-SEP-2005 (TREMBLrel. 31, Created)
DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Predicted protein involved in purine metabolism.
GN OrderedLocustNames=NT10757;
OS Haemophilus influenzae (strain 86-028NP).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=281310;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
RA Harrison A., Dyer D.W., Gallaspy A., Ray W.C., Mungur R., Carson M.B.,
RA Zhong H., Gibson J., Gibson M., Johnson L.S., Lewis L., Bakaletz L.O.,
RA Munson R.S. Jr.;
RA "Genomic sequence of an otitis media isolate of nonypeable
RT Haemophilus influenzae: comparative study with H. influenzae serotype
RT d, strain KW20."
RL J. Bacteriol. 187:4627-4636 (2005).
DR EMBL: CP000057; AAX87668.1; -; Genomic_DNA.
DR InterPro: IPR007451; DUF489.
DR Pfam: PF04356; DUF489; 1.
KM Complete proteome.
SQ SEQUENCE 205 AA; 23221 MW; AB9F6AC75D446DF2 CRC64;

Query Match
Best Local Similarity 79.3%; Score 92; DB 2; Length 205;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NYVDITLALAGVCOAKLVQCPA 23

```

Db 3 NYHDIYALAGVCO\$AKLWHOLA 25

Search completed: January 23, 2006, 09:31:06
Job time : 61.7 secs

This Page Blank (uspo:0)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-2

Perfect score: 116

Sequence: 1 NYVDITLALAGVCAKLVOQFA 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	88.8	208	2 AH0199	conserved hypothe
2	103	88.8	213	2 S19211	ycfc protein - Esc
3	103	88.8	213	2 D90829	hypothetical prote
4	103	88.8	213	2 B85687	hypothetical prote
5	100	86.2	215	2 AB0647	conserved hypothe
6	92	79.3	205	2 I64155	hypothetical prote
7	72	62.1	205	2 B82337	conserved hypothe
8	71	61.2	211	2 D84960	hypothetical prote
9	50	43.1	384	2 I57957	neurokinin 2 recep
10	50	43.1	384	2 S20303	neurokinin 2 recep
11	50	43.1	390	2 A36727	neurokinin 2 recep
12	49	42.2	276	2 C86526	hypothetical prote
13	49	42.2	332	2 E72097	hypothetical prote
14	49	42.2	384	1 S00516	neurokinin 2 recep
15	49	42.2	358	1 JQ1059	neurokinin 2 recep
16	49	42.2	402	2 I56595	early E3 18.5K gly
17	46	39.7	159	1 O6ADE	conserved hypothe
18	46	39.7	204	2 B82680	nucleoside prote
19	45.5	38.8	243	1 VHVUPT	heat shock protein
20	45	38.8	142	2 AP0961	heat shock protein
21	45	38.8	144	2 B91207	heat shock protein
22	45	38.8	144	2 D65170	heat shock protein
23	45	38.8	144	2 D86053	heat shock protein
24	45	38.8	847	2 G95843	conserved hypothe
25	44	37.9	165	2 AB0851	hypothetical prote
26	44	37.9	165	2 S70217	slpB protein - Sal
27	44	37.9	165	2 A57357	slpB protein - Sal
28	44	37.9	188	2 JC1065	coat protein - bee
29	44	37.9	259	2 S14731	bacteriorhodopsin

30	44	37.9	260	2 A34178	bacteriorhodopsin
31	44	37.9	280	2 A10326	cel operon repres
32	44	37.9	336	2 AG0105	sugar transport sy
33	44	37.9	357	2 T26668	hypothetical prote
34	43	37.1	457	2 A95940	probable integrat
35	43	37.1	464	2 B87573	pyrazinamide/nic
36	43	37.1	538	2 A41042	microbial metallo
37	42	36.2	84	2 C86579	CT666 hypothe
38	42	36.2	84	2 F72046	conserved hypothe
39	42	36.2	154	2 AC0496	heat shock protein
40	42	36.2	204	2 C47686	bacteriorhodopsin
41	42	36.2	273	2 AC3277	transposase BME102
42	42	36.2	295	2 D82443	transcription regu
43	42	36.2	335	2 S70671	lipopolysaccharide
44	42	36.2	407	2 S20304	neurokinin 1 recep
45	42	36.2	407	2 A34357	neurokinin 1 recep

ALIGNMENTS

RESULT 1
AH0199
conserved hypothetical protein YP01637 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence revision 02-Nov-2001 #ext_change 09-Jul-2004
C/Accession: AH0199
R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
H. M.; Rutherford, K.; Skelton, J.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AH0199
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-208 <KUR>
A/Cross-references: UNIPROT:Q8ZF06; UNIPARC:UPI00000CD828; GB:AL590842; PIDN:CAC90459.1
C/Genetics:
A/Gene: YP01637

Query Match 88.8%; Score 103; DB 2; Length 208;
Best Local Similarity 82.6%; Pred. No. 2.1e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYVDITLALAGVCAKLVOQFA 23
DB 4 NYVDITLALAGVCAKLVOQFA 26

RESULT 2
S19211
ycfc protein - Baccharichia coli (strain K-12)
C/Species: Baccharichia coli
C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #ext_change 09-Jul-2004
C/Accession: S19211; A64858
R/Green, S.M.; Drabble, W.T.
submitted to the EMBL Data Library, May 1991
A/Description: Molecular analysis of the purB-phop region of Baccharichia coli K12.
A/Reference number: S19210
A/Accession: S19211
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-213 <GRB>
A/Cross-references: UNIPROT:P25746; UNIPARC:UPI000013A673; EMBL:X59307; MID:942582; PID
A/Experimental source: UNIPROT: P25746; UNIPARC: UPI000013A673; EMBL: X59307; MID: 942582; PID
R/Blaetner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C

.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Baccharichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: A64858
A/Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-213 <BLAT>
A:Cross-references: UNIPARC:UPI000013A673; GB:AB000213; GB:U00096; NID:g1787371; PIDN:AA
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ycfC
A:Start codon: CTG

Query Match 88.8%; Score 103; DB 2; Length 213;
Best Local Similarity 82.6%; Pred. No. 2.1e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 NYDITLALAGVCOAALVQQA 23
DB 4 NYDITLALAGVCOAALVQQA 26

RESULT 3
D90829
hypotheetical protein EC81604 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D90829
R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gagawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Rep. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; PMID:1156231; PMID:11258796
A:Accession: D90829
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <HAV>
A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:BA000007; PIDN:BA035027.1;
A:Experimental source: strain O157:H7, substrain RMD 050952
C:Genetics:
A:Gene: EC81604

Query Match 88.8%; Score 103; DB 2; Length 213;
Best Local Similarity 82.6%; Pred. No. 2.1e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 NYDITLALAGVCOAALVQQA 23
DB 4 NYDITLALAGVCOAALVQQA 26

RESULT 4
B85687
hypotheetical protein ycfC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85687
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grochocinski, E.J.; Davis, N.W.; Lam, A.; Dimmler, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; PMID:12074935; PMID:11206551
A:Accession: B85687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:AB005174; NID:g12514786; F
C:Genetics:
A:Gene: ycfC

Query Match 88.8%; Score 103; DB 2; Length 213;
Best Local Similarity 82.6%; Pred. No. 2.1e-09;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 NYDITLALAGVCOAALVQQA 23
DB 4 NYDITLALAGVCOAALVQQA 26

RESULT 5
AB0647
conserved hypotheetical protein STY1273 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typh
A>Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0647
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moul, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AB0647
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-215 <PAR>
A:Cross-references: UNIPARC:UPI0000059F77; GB:AL513382; PIDN:CAD08357.1; PID:g16502402;
C:Genetics:
A:Gene: STY1273

Query Match 86.2%; Score 100; DB 2; Length 215;
Best Local Similarity 78.3%; Pred. No. 6.8e-09;
Matches 18; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

1 NYDITLALAGVCOAALVQQA 23
DB 6 NYDITLALAGVCOAALVQQA 28

RESULT 6
164155
hypotheetical protein H10638 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: 164155
R:Rifelschman, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fuhremann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; PMID:95350630; PMID:7542800
A:Accession: 164155
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-205 <TIGR>
A:Cross-references: UNIPROT:P44796; UNIPARC:UPI000013A674; GB:U32747; GB:LA2023; NID:g15
A>Note: best homolog was a hypotheetical protein from Escherichia coli

Query Match 79.3%; Score 92; DB 2; Length 205;
Best Local Similarity 78.3%; Pred. No. 1.4e-07;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

1 NYDITLALAGVCOAALVQQA 23
DB 3 NYDITLALAGVCOAALVQQA 25

RESULT 7
B82237
conserved hypotheetical protein VC1127 [imported] - Vibrio cholerae (strain N16961 serogr.
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: B82237
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chadson, D.; Brimicombe, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: B82237

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-205 <HEI>

A/Cross-references: UNIPROT:Q9KX9; UNIPARC:UPI00000C2EB4; GB:AE004193; GB:AE003852; NID

A/Experimental source: serogroup O1, strain M16961; biotype E1 Tor

C/Genetics:

A/Gene: VC1127

A/Map position: 1

Query Match 62.1%; Score 72; DB 2; Length 205;

Best Local Similarity 66.7%; Pred. No. 0.00029;

Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 NYDITLALAGVCOAAKVQCPA 23

DB 6 YDRTIAPGICQAVLVQCPA 26

RESULT 8

hypothetical protein [imported] - Buchnera sp. (strain APS)

C/Species: Buchnera sp.

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001

C/Accession: D84960

R/Shigemitsu, S.; Matsumoto, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A/Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A

A/Reference number: A84930; MUID:20445173; PMID:10930307

A/Accession: D84960

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1211 <STO>

A/Cross-references: UNIPARC:UPI000005E4F; GB:AP000398; GSPDB:GN00144

A/Experimental source: strain APS

C/Genetics:

A/Gene: ycfC; BU262

Query Match 61.2%; Score 71; DB 2; Length 211;

Best Local Similarity 73.7%; Pred. No. 0.00044;

Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 ITTLALAGVCOAAKLVQCPA 23

DB 7 ITTLALAGVCOAAKLVQCPA 25

RESULT 9

neurokinin 2 receptor - hamster

N/Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)

C/Species: Cricetinae gen. sp. (hamster)

C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 19-May-2000

C/Accession: I57957

R/Aharony, D.; Little, J.; Thomas, C.; Powell, S.; Berry, D.; Graham, A.

Mol. Pharmacol. 45, 9-19, 1994

A/Title: Isolation and pharmacological characterization of a hamster urinary bladder neu

A/Reference number: I57957; MUID:94134065; PMID:8302285

A/Accession: I57957

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-384 <RSS>

A/Cross-references: UNIPARC:UPI00001301F1; GB:S68899; NID:G545230; PIDN:AC60680.1; PID

C/Superfamily: neurokinin 1 receptor

Query Match 43.1%; Score 50; DB 2; Length 384;

Best Local Similarity 62.5%; Pred. No. 2.4;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAA 16

DB 69 NYPIINALADLCMAA 84

RESULT 10

neurokinin 2 receptor - mouse

N/Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)

C/Species: Mus musculus (house mouse)

C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C/Accession: S20303

R/Sundelin, J.B.; Provvedini, D.M.; Wahlestedt, C.R.; Laurell, H.; Pohl, J.S.; Peterson,

Eur. J. Biochem. 203, 625-631, 1992

A/Title: Molecular cloning of the murine substance K and substance P receptor genes.

A/Reference number: S20303; MUID:92137253; PMID:1370937

A/Accession: S20303

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-384 <STO>

A/Cross-references: UNIPROT:P30549; UNIPARC:UPI000029851; GB:X62933; NID:G54204; PIDN:

C/Superfamily: neurokinin 1 receptor

C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 43.1%; Score 50; DB 2; Length 384;

Best Local Similarity 62.5%; Pred. No. 2.4;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAA 16

DB 69 NYPIINALADLCMAA 84

RESULT 11

neurokinin 2 receptor - rat

N/Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)

C/Species: Rattus norvegicus (Norway rat)

C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004

C/Accession: A36737

R/Sasal, Y.; Nakanishi, S.

Biochem. Biophys. Res. Commun. 165, 695-702, 1989

A/Title: Molecular characterization of rat substance K receptor and its mRNA.

A/Reference number: A36737; MUID:90088481; PMID:2480781

A/Accession: A36737

A/Molecule type: mRNA

A/Residues: 1-390 <SAS>

A/Cross-references: UNIPROT:P16610; UNIPARC:UPI00001301F3; GB:M1838; NID:G206986; PIDN

C/Superfamily: neurokinin 1 receptor

C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 43.1%; Score 50; DB 2; Length 390;

Best Local Similarity 62.5%; Pred. No. 2.4;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAA 16

DB 69 NYPIINALADLCMAA 84

RESULT 12

hypothetical protein CP10283 [imported] - Chlamydia pneumoniae (strain J138)

C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: C86526

R/Shirai, M.; Hisekawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I

Nucleic Acids Res. 28, 2311-2314, 2000

A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A/Reference number: A86491; MUID:20330349; PMID:10871162

A/Accession: C86526

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-276 <STO>

A/Cross-references: UNIPROT:Q9JSG6; UNIPARC:UPI00000CCD08; GB:BA000008; NID:G8978657; P

A/Experimental source: strain J138

C:Genetics:
A:Gene: CPJ0283
C:Superfamily: Chlamydia pneumoniae hypothetical protein CPn0283

Query Match 42.2%; Score 49; DB 2; Length 276;
Best Local Similarity 42.3%; Pred. No. 2.6;
Matches 11; Conservative 4; Mismatches 3; Indels 8; Gaps 1;

Oy 5 ITTALAGVC-----QAKVQOF 22
Db 35 INLLSGCLRRSHDRSQAVKIQOF 60

RESULT 13

hypothetical protein CPn0283 - Chlamydia pneumoniae (strain CWL029)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: E72097

R:Kaihan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: E72097

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-322 <ARN>

A:Cross-references: UNIPROT:092805; UNIPARC:UPI00000C11C8; GB:AE001613; GB:AE001363; NID

A:Experimental source: strain CWL029

C:Genetics:

A:Gene: CPn0283

C:Superfamily: Chlamydia pneumoniae hypothetical protein CPn0283

Query Match 42.2%; Score 49; DB 2; Length 322;
Best Local Similarity 42.3%; Pred. No. 3;
Matches 11; Conservative 4; Mismatches 3; Indels 8; Gaps 1;

Oy 5 ITTALAGVC-----QAKVQOF 22
Db 81 INLLSGCLRRSHDRSQAVKIQOF 106

RESULT 14

S00516

neurokinin 2 receptor - bovine

N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: S00516

R:Maqu, Y.; Nakayama, K.; Tanaki, H.; Harada, Y.; Kuno, M.; Nakanishi, S.

Nature 329, 836-838, 1987

A:Title: cDNA cloning of bovine substance K receptor through oocyte expression system.

A:Reference number: S00516; MUID:88039072; PMID:2823146

A:Accession: S00516

A:Molecule type: mRNA

A:Residues: 1-384 <MA>

A:Cross-references: UNIPROT:P05363; UNIPARC:UPI00001301BE; EMBL:X06295; NID:g746; PIDN:C

C:Superfamily: neurokinin 1 receptor

C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane pr

F:33-59/Domain: transmembrane #status predicted <TM2>

F:70-93/Domain: transmembrane #status predicted <TM2>

F:111-129/Domain: transmembrane #status predicted <TM3>

F:150-169/Domain: transmembrane #status predicted <TM4>

F:193-222/Domain: transmembrane #status predicted <TM5>

F:252-275/Domain: transmembrane #status predicted <TM6>

F:288-310/Domain: transmembrane #status predicted <TM7>

F:11,19/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:106-181/Disulfide bonds: #status predicted

Query Match 42.2%; Score 49; DB 1; Length 384;

Best Local Similarity 56.2%; Pred. No. 3.5;

Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 NYDITTLAGVCOAA 16
Db 69 NYFIVNLALADLCMAA 84

RESULT 15

neurokinin 2 receptor - human

N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)

C:Species: Homo sapiens (man)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C:Accession: J01059; J05075; A23658; A61224

R:Graham, A.; Hopkins, B.; Powell, S.J.; Danks, P.; Briggs, I.

Biochem. Biophys. Res. Commun. 177, 8-16, 1991

A:Title: Isolation and characterization of the human lung NK-2 receptor gene using rapid

A:Reference number: J01059; MUID:91253441; PMID:1710456

A:Accession: J01059

A:Molecule type: DNA

A:Residues: 1-398 <GRAL>

A:Cross-references: UNIPROT:P21452; UNIPARC:UPI000016ADE7; GB:M75105; NID:g189219; PIDN:

A:Accession: J05075

A:Molecule type: mRNA

A:Residues: 1-398 <GRAL2>

A:Cross-references: UNIPARC:UPI000016ADE7; GB:M75105; NID:g189219; PIDN:AAB05897.1; PID:

A:Experimental source: lung

A>Note: 23-ile and 375 His were also found

R:Gerard, N.P.; Eddy Jr., R.L.; Shows, T.B.; Gerard, C.

J. Biol. Chem. 265, 20455-20462, 1990

A:Title: The human neurokinin A (substance K) receptor. Molecular cloning of the gene, c

A:Reference number: A23658; MUID:91056095; PMID:2173708

A:Accession: A23658

A:Molecule type: DNA

A:Residues: 1-22, 'I', 24-240, 'L', 242-398 <GER>

A:Cross-references: UNIPARC:UPI000017351F; GB:M60264; GB:J05680; NID:g189140

A>Note: the authors translated the codon GGA for residue 317 as Glu

R:Xiris, R.M.; South, V.; Saltzman, A.; Felder, S.; Ricca, G.A.; Jaffe, M.; Huebner, K.; K

Cell Growth Differ. 2, 15-22, 1991

A:Title: Cloning and expression of the human substance K receptor and analysis of its ro

A:Reference number: A61224; MUID:911175483; PMID:1848773

A:Accession: A61224

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-22, 'I', 24-293, 'F', 294-374, 'H', 376-398 <KRI>

A:Cross-references: UNIPARC:UPI0000173520

C:Comment: The endogenous ligand of this receptor is neurokinin 2 (substance K or neurok

C:Genetics:

A:Gene: GDB:TAC2R

A:Cross-references: GDB:126367; OMIM:162321

A:Map position: 10q11-10q21

A:introns: 131/2; 247/3; 313/2

C:Superfamily: neurokinin 1 receptor

C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane pr

F:33-59/Domain: transmembrane #status predicted <TM2>

F:70-93/Domain: transmembrane #status predicted <TM2>

F:111-129/Domain: transmembrane #status predicted <TM3>

F:150-169/Domain: transmembrane #status predicted <TM4>

F:193-222/Domain: transmembrane #status predicted <TM5>

F:252-275/Domain: transmembrane #status predicted <TM6>

F:288-310/Domain: transmembrane #status predicted <TM7>

F:11,19/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:106-181/Disulfide bonds: #status predicted

Query Match 42.2%; Score 49; DB 1; Length 398;

Best Local Similarity 56.2%; Pred. No. 3.7;

Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 NYDITTLAGVCOAA 16
Db 69 NYFIVNLALADLCMAA 84

Search completed: January 23, 2006, 09:32:48
Job time : 10.1 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 Seconds
(without alignments)
114.551 Million cell updates/sec

Title: US-10-501-838a-2

Perfect score: 116
Sequence: 1 NYDITLALAGVCOAKLVQQA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/1aa/5 COMB .pep:.*
2: /cgn2_6/prodata/1/1aa/6 COMB .pep:.*
3: /cgn2_6/prodata/1/1aa/H COMB .pep:.*
4: /cgn2_6/prodata/1/1aa/PCTUS COMB .pep:.*
5: /cgn2_6/prodata/1/1aa/RR COMB .pep:.*
6: /cgn2_6/prodata/1/1aa/backfillseq.pep:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	92.2	217	2 US-09-489-039A-7789	Sequence 7789, Ap
2	85	73.3	216	2 US-09-543-681A-7993	Sequence 7993, Ap
3	50	43.1	369	1 US-07-937-609-19	Sequence 19, Appl
4	50	43.1	369	2 US-08-029-170-19	Sequence 19, Appl
5	50	43.1	369	2 US-09-443-745-19	Sequence 19, Appl
6	49	42.2	94	2 US-09-198-452A-1181	Sequence 1181, Ap
7	49	42.2	350	2 US-09-438-185A-285	Sequence 285, Appl
8	49	42.2	384	1 US-08-103-170-10	Sequence 10, Appl
9	49	42.2	387	1 US-08-196-989B-14	Sequence 14, Appl
10	49	42.2	387	1 US-08-760-936-14	Sequence 14, Appl
11	49	42.2	387	2 US-09-225-024-14	Sequence 14, Appl
12	49	42.2	398	2 US-09-826-509-527	Sequence 527, Appl
13	46	39.7	433	2 US-09-270-767-43172	Sequence 43172, A
14	46	39.7	433	2 US-09-270-767-43172	Sequence 43172, A
15	45.5	39.2	440	2 US-09-605-703B-1288	Sequence 1288, Ap
16	45.5	39.2	522	2 US-09-605-703B-1288	Sequence 1288, Ap
17	45	38.8	169	2 US-09-489-039A-10685	Sequence 10685, A
18	44	37.9	309	2 US-09-902-540-12828	Sequence 12828, A
19	44	37.9	400	2 US-09-543-681A-4587	Sequence 4587, Ap
20	44	37.9	462	1 US-08-471-033-20	Sequence 20, Appl
21	44	37.9	462	1 US-08-471-033-20	Sequence 20, Appl
22	44	37.9	462	1 US-08-463-483A-20	Sequence 20, Appl
23	44	37.9	462	1 US-08-471-046A-20	Sequence 20, Appl
24	44	37.9	462	1 US-08-470-566B-20	Sequence 20, Appl
25	44	37.9	462	1 US-08-469-334-20	Sequence 20, Appl
26	44	37.9	462	2 US-09-300-529-20	Sequence 20, Appl
27	43	37.1	182	2 US-09-902-540-13669	Sequence 13669, A

ALIGNMENTS

28	43	37.1	335	2	US-09-252-991A-26794	Sequence 26794, A
29	42	36.2	60	2	US-09-248-796A-23018	Sequence 23018, A
30	42	36.2	90	2	US-09-124-671-25	Sequence 25, Appl
31	42	36.2	91	2	US-09-198-452A-753	Sequence 753, Appl
32	42	36.2	91	2	US-09-438-185A-712	Sequence 712, Appl
33	42	36.2	109	2	US-09-124-671-23	Sequence 23, Appl
34	42	36.2	109	2	US-09-124-671-23	Sequence 23, Appl
35	42	36.2	109	2	US-09-124-671-29	Sequence 29, Appl
36	42	36.2	372	1	US-07-937-609-20	Sequence 20, Appl
37	42	36.2	372	2	US-08-029-170-20	Sequence 20, Appl
38	42	36.2	372	2	US-09-443-745-20	Sequence 20, Appl
39	42	36.2	398	2	US-09-543-681A-7046	Sequence 7046, Ap
40	42	36.2	407	4	PCT-US92-06532-7	Sequence 7, Appl
41	41	35.3	114	2	US-09-902-540-10640	Sequence 10640, A
42	41	35.3	231	2	US-09-830-433A-32	Sequence 32, Appl
43	41	35.3	407	1	US-08-390-000A-6	Sequence 6, Appl
44	41	35.3	411	2	US-09-328-352-6719	Sequence 6719, Ap
45	41	35.3	741	2	US-09-252-991A-27062	Sequence 27062, A

RESULT 1
US-09-489-039A-7789
Sequence 7789, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489, 039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7789
LENGTH: 217
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match 92.2%; Score 107; DB 2; Length 217;
Best Local Similarity 91.3%; Pred. No. 2, 6e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAKLVQQA 23
DB 8 NYDITLALAGVCOAKLVQQA 30

RESULT 2
US-09-543-681A-7993
Sequence 7993, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543, 681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7993
LENGTH: 216
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7993
Query Match 73.3%; Score 85; DB 2; Length 216;

Best Local Similarity 69.6%; Pred. No. 8.6e-06;
Matches 16; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYDITLALAGVCOAKLVQQA 23
Db 11 DFDITLALAGVCOAKLVQQA 33

RESULT 3

US-07-937-609-19
Sequence 19, Application US/07937609
Patent No. 5319073

GENERAL INFORMATION:

APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-07-937-609-19

Query Match 43.1%; Score 50; DB 1; Length 369;
Best Local Similarity 62.5%; Pred. No. 6.4;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYDITLALAGVCOAA 16
Db 69 NYFINLALADLCWAA 84

RESULT 4

US-08-029-170-19
Sequence 19, Application US/08029170
Patent No. 6169173
GENERAL INFORMATION:

APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-08-029-170-19

Query Match 43.1%; Score 50; DB 2; Length 369;
Best Local Similarity 62.5%; Pred. No. 6.4;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYDITLALAGVCOAA 16
Db 69 NYFINLALADLCWAA 84

RESULT 5

US-09-443-745-19
Sequence 19, Application US/09443745
Patent No. 6706493

GENERAL INFORMATION:

APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA

ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance X receptor
US-09-443-745-19

Query Match 43.1%; Score 50; DB 2; Length 369;
Best Local Similarity 62.5%; Pred. No. 6.4;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGVQQA 16
Db 69 NYPIMLALDLCPAA 84

RESULT 6
US-09-198-452A-1181
Sequence 1181, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6649
SEQ ID NO 1181
LENGTH: 94
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1181

Query Match 42.2%; Score 49; DB 2; Length 94;
Best Local Similarity 42.3%; Pred. No. 2;
Matches 11; Conservative 4; Mismatches 3; Indels 8; Gaps 1;

QY 5 ITLALAGVC-----QAAKLVQGF 22
Db 35 INLLSGCLRSSHDSQAVKTIQGF 60

RESULT 7
US-09-438-185A-285
Sequence 285, Application US/09438185A
Patent No. 6822071
GENERAL INFORMATION:
APPLICANT: Stephens, Richard
APPLICANT: Mitchell, Wayne
APPLICANT: Kalman, Sue
APPLICANT: Davis, Ronald
TITLE OF INVENTION: The Regents of the University of California
FILE REFERENCE: 018941-000411US
CURRENT APPLICATION NUMBER: US/09/438,185A
FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/128,606
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 1074
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 285
LENGTH: 350
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
FEATURE:
OTHER INFORMATION: CPM0283
US-09-438-185A-285

Query Match 42.2%; Score 49; DB 2; Length 350;
Best Local Similarity 42.3%; Pred. No. 8.7;
Matches 11; Conservative 4; Mismatches 3; Indels 8; Gaps 1;

QY 5 ITLALAGVC-----QAAKLVQGF 22
Db 109 INLLSGCLRSSHDSQAVKTIQGF 134

RESULT 8
US-08-103-170-10
Sequence 10, Application US/08103170
Patent No. 5865824
GENERAL INFORMATION:
APPLICANT: Yamada, Tadataka
APPLICANT: Gantz, Ira
TITLE OF INVENTION: Recombinant Genomic Clones Encoding
TITLE OF INVENTION: Histamine H1, H2, and H3 Receptors, Methods For Production
TITLE OF INVENTION: Thereof, and Proteins Encoded Therefrom
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/633,060
FILING DATE: 24-DEC-1990

ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-017-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 521-4500
TELEFAX: (703) 486-2347
TELEX: 24885 OPAT UR
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bos taurus
US-08-103-170-10

Query Match 42.2%; Score 49; DB 1; Length 384;
Best Local Similarity 56.2%; Pred. No. 9.7;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYDITLALAGVCOAA 16
Db 69 NYFVNLALADLCMAA 84

RESULT 9
US-08-196-989B-14
Sequence 14, Application US/08196989B
Patent No. 5585476
GENERAL INFORMATION:
APPLICANT: MacLennan, A. John
TITLE OF INVENTION: Molecular Cloning and Expression of
TITLE OF INVENTION: G-Protein Coupled Receptors
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSER: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/196,989B
FILING DATE: 15-FEB-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Lloyd, Jeff
REGISTRATION NUMBER: 35,589
REFERENCE/DOCKET NUMBER: MAC-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-196-989B-14

Query Match 42.2%; Score 49; DB 1; Length 387;
Best Local Similarity 56.2%; Pred. No. 9.7;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
Qy 1 NYDITLALAGVCOAA 16

Db 69 NYFVNLALADLCMAA 84

RESULT 10
US-08-760-936-14
Sequence 14, Application US/08760936
Patent No. 5856443
GENERAL INFORMATION:
APPLICANT: MacLennan, A. John
TITLE OF INVENTION: Molecular Cloning and Expression of
TITLE OF INVENTION: G-Protein Coupled Receptors
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,936
FILING DATE: December 6, 1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: MAC-100C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-760-936-14

Query Match 42.2%; Score 49; DB 1; Length 387;
Best Local Similarity 56.2%; Pred. No. 9.7;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYDITLALAGVCOAA 16
Db 69 NYFVNLALADLCMAA 84

RESULT 11
US-09-225-024-14
Sequence 14, Application US/09225024
Patent No. 6518414
GENERAL INFORMATION:
APPLICANT: MacLennan, A. John
TITLE OF INVENTION: Molecular Cloning and Expression of
TITLE OF INVENTION: G-Protein Coupled Receptors
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/225,024
FILING DATE: 04-JAN-1999
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/760,936
FILING DATE: 6-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/196,989
FILING DATE: 15-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: MAC-100C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-225-024-14

Query Match 42.2%; Score 49; DB 2; Length 387;
Best Local Similarity 56.2%; Pred. No. 9.7;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAA 16
||:|||||:
DB 69 NYFIVNALADLCMAA 84

RESULT 12
US-09-826-509-527
Sequence 527, Application US/09826509
Patent No. 6806054
GENERAL INFORMATION:
APPLICANT: Lehmann-Brulsma, Karin
APPLICANT: Law, Chen W.
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: Patentin Version 2.1
SEQ ID NO 527
LENGTH: 398
TYPE: PRT
ORGANISM: Homo sapiens
US-09-826-509-527

Query Match 42.2%; Score 49; DB 2; Length 398;
Best Local Similarity 56.2%; Pred. No. 10;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAA 16
||:|||||:
DB 69 NYFIVNALADLCMAA 84

RESULT 13
US-09-270-767-58510

Sequence 58510, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 58510
LENGTH: 138
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-58510

Query Match 39.7%; Score 46; DB 2; Length 138;
Best Local Similarity 52.9%; Pred. No. 9.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAAK 17
||:|||||:
DB 24 NYADIEINALGLADKAR 40

RESULT 14
US-09-270-767-43172
Sequence 43172, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 43172
LENGTH: 433
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43172

Query Match 39.7%; Score 46; DB 2; Length 433;
Best Local Similarity 52.9%; Pred. No. 33;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYDITLALAGVCOAAK 17
||:|||||:
DB 319 NYADIRIALGLADKAR 335

RESULT 15
US-09-605-703B-1288
Sequence 1288, Application US/09605703B
Patent No. 6962989
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberbauer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
FILE REFERENCE: BGI-129CP
CURRENT APPLICATION NUMBER: US/09/605,703B
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/142,764
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 60/152,318
PRIOR FILING DATE: 1999-09-03

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:32:56 ; Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838a-2

Perfect score: 116
Sequence: 1 NYDITLALAGVCAKLVQCPA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.5	39.2	522	7	US-11-055-822-1100 Sequence 1100, App
2	43	37.1	42	6	US-10-512-295A-3 Sequence 3, Appl
3	40	34.5	63	6	US-10-467-657-6638 Sequence 6638, App
4	39.5	34.1	447	6	US-10-858-730-219 Sequence 219, App
5	39	33.6	1015	6	US-10-957-569-51 Sequence 51, Appl
6	38.5	33.2	341	6	US-10-467-657-154 Sequence 154, App
7	38.5	33.2	341	6	US-10-467-657-5822 Sequence 5822, App
8	38.5	33.2	341	6	US-10-467-657-7476 Sequence 7476, App
9	38	32.8	282	6	US-10-467-657-4418 Sequence 4418, App
10	37.5	32.3	311	7	US-11-055-822-94 Sequence 94, Appl
11	37.5	32.3	447	7	US-11-055-822-94 Sequence 94, Appl
12	37	31.9	54	6	US-10-467-657-1542 Sequence 1542, App
13	37	31.9	268	7	US-11-055-822-926 Sequence 926, App
14	37	31.9	330	7	US-11-055-822-924 Sequence 924, App
15	37	31.9	696	7	US-11-052-554A-100 Sequence 100, App
16	37	31.9	696	7	US-11-052-554A-101 Sequence 101, App
17	37	31.9	787	6	US-10-467-657-2832 Sequence 2832, App
18	37	31.9	8695	7	US-11-205-109-15 Sequence 15, Appl
19	36.5	31.5	162	6	US-10-467-657-2510 Sequence 2510, App
20	36.5	31.5	525	7	US-11-112-882-27 Sequence 27, Appl
21	36	31.0	203	6	US-10-453-372-890 Sequence 890, App
22	36	31.0	367	6	US-10-821-234-1569 Sequence 1569, App
23	36	31.0	473	7	US-11-069-642-5 Sequence 5, Appl
24	36	31.0	1268	7	US-11-052-554A-1 Sequence 1, Appl
25	36	31.0	1732	6	US-10-055-877-147 Sequence 147, App

26	36	31.0	3002	6	US-10-821-234-916 Sequence 916, App
27	35	30.2	45	6	US-10-467-657-8968 Sequence 8968, App
28	35	30.2	138	7	US-11-051-453-56 Sequence 56, Appl
29	35	30.2	185	6	US-10-453-372-896 Sequence 896, App
30	35	30.2	194	6	US-10-453-372-896 Sequence 894, App
31	35	30.2	205	6	US-10-793-626-1828 Sequence 1828, App
32	35	30.2	218	6	US-10-453-372-892 Sequence 892, App
33	35	30.2	263	6	US-10-793-626-1418 Sequence 1418, App
34	35	30.2	270	6	US-10-467-657-3022 Sequence 3022, App
35	35	30.2	305	7	US-11-000-463-391 Sequence 391, App
36	35	30.2	305	7	US-11-000-463-391 Sequence 863, App
37	35	30.2	474	7	US-11-000-463-284 Sequence 284, App
38	35	30.2	480	7	US-11-136-244-9 Sequence 9, Appl
39	35	30.2	491	6	US-10-763-712A-4 Sequence 4, Appl
40	35	30.2	519	6	US-11-183-914-9 Sequence 9, Appl
41	35	30.2	519	6	US-10-999-886-3 Sequence 3, Appl
42	35	30.2	619	7	US-11-136-244-3 Sequence 3, Appl
43	35	30.2	640	6	US-10-999-886-4 Sequence 4, Appl
44	35	30.2	640	7	US-11-136-244-4 Sequence 4, Appl
45	35	30.2	750	7	US-11-070-627-6 Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-11-055-822-1100
; Sequence 1100, Application US/11055822
; Publication No. US20050260707A1
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Kroeger, Burkhard
APPLICANT: Schroeder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauser, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
TITLE OR INVENTION: METABOLIC PATHWAY PROTEINS
FILE REFERENCE: BGI-121CPCN
CURRENT APPLICATION NUMBER: US/11/055,822
PRIOR FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: 09/606,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142,101
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/148,613
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/187,970
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: DE 19930476.9
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19931415.2
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931418.7
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
Remaining Prior Application data removed - See File Wrapper or PALM.
SEQ ID NO 1100
LENGTH: 522
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-11-055-822-1100

Query Match 39.2%; Score 45.5; DB 7; Length 522;
Best Local Similarity 44.1%; Pred. No. 3.6;
Matches 15; Conservative 1; Mismatches 7; Indels 11; Gaps 1;
CY 1 NYDITLALAGVCAKLVQCPA 23

Db 289 NPYDIKEEVGLTEQLTLASAGVFYAAEARQAF 322

```

RESULT 2
US-10-512-295A-3
Sequence 3, Application US/10512295A
Publication No. US20050245727A1
GENERAL INFORMATION:
APPLICANT: Baltzer, Lars
APPLICANT: Dolphin, Gunnar
APPLICANT: Liedberg, Bo
APPLICANT: Lundstrom, Ingemar
TITLE OF INVENTION: NOVEL POLYPEPTIDE SCAFFOLDS AND USE THEREOF
FILE REFERENCE: 5846.181USNO
CURRENT APPLICATION NUMBER: US/10/512.295A
CURRENT FILING DATE: 2004-09-24
PRIOR APPLICATION NUMBER: PCT/SE03/00507
PRIOR FILING DATE: 2003-03-26
PRIOR APPLICATION NUMBER: SE 0200968-6
PRIOR FILING DATE: 2002-03-26
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.3
SEQ ID NO 3
LENGTH: 42
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Peptide KE3
US-10-512-295A-3

```

Query Match	37.1%	Score 43;	DB 6;	Length 42;
Best Local Similarity	52.9%	Pred. No. 0.66;		
Matches	9;	Conservative	2;	Mismatches 6;
				Indels 0;
				Gaps 0;

```
Qy      7 LALAGVCQAQKLVQQFA 23
        ||| | | | : | |
Db     16 LARGPCDAAQLAEQLA 32
```

```

RESULT 3
US-10-467-657-6638
Sequence 6638, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHITRON SpA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqMan99, version 1.04
SEQ ID NO 6638
LENGTH: 63
TYPE: PR1
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6638

```

Query Match	34.5%	Score 40;	DB 6;	Length 63;
Best Local Similarity	53.3%	Pred. No. 3.2;		
Matches	8;	Conservative	2;	Mismatches 5;
				Indels 0;
				Gaps 0;

```
QY      2 YYDITLALAGVCQAA 16
          ||:|:|
Db      46 YYNITVASTVKCQTA 60
```

```

RESULT 4
US-10-858-730-219
; Sequence 219, Application US/10658730
; Publication No. US20050255566A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858, 730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-10
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219
; LENGTH: 447
; TYPE: PRF
; ORGANISM: Coryne-bacterium glutamicum
US-10-858-730-219

```

Query Match	34.1%	Score 39.5;	DB 6;	length 447;
Best Local Similarity	36.7%	Pred. NO. 30;		
Matches 11; Conservative	4;	Mismatches 4;	Indels 11;	Gaps 1;

```

QY      1 NYDITL-----ALAGVCGAKLV 13
          |||: | :| :| :|
DB      9 NYDMLLKRNAGEPEFHQVAVAEVLSELTIV 38

```

```

RESULT 5
US-10-957-569-51
; Sequence 51, Application US/10957569
; Publication No. US20050246785A1
; GENERAL INFORMATION:
; APPLICANT: COOK, Zhihong et al.
; TITLE OF INVENTION: PROMOTER, PROMOTER CONTROL ELEMENTS, AND COMBINATIONS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 2750-1577PUS3
; CURRENT APPLICATION NUMBER: US/10/957,569
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 10/950,321
; PRIOR FILING DATE: 2004-09-23
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-957-569-51

```

Query Match	33.6%	Score 39;	DB 6;	Length 105;
Best Local Similarity	43.8%	Pred. No. 87;		
Matches	7;	Conservative	3;	Mismatches 6;
				Indels 0;
				Gaps 0

QY 4 DITLALAGVCCQAAKLV 19
: : : : : :
Db 172 DMTLMIIGVCAFSVLI 187

RESULT 6
US-10-467-657-154

```

; Sequence 154, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 154
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-154

Query Match      33.2%; Score 38.5; DB 6; Length 341;
Best Local Similarity 45.5%; Pred. No. 33;
Matches 10; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY      3 YDITLALA-----GVCQAAKLV 19
DB      148 YDVTLTMANAKDYGVAQERKRV 169

RESULT 7
US-10-467-657-5822
; Sequence 5822, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5822
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5822

Query Match      33.2%; Score 38.5; DB 6; Length 341;
Best Local Similarity 45.5%; Pred. No. 33;
Matches 10; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY      3 YDITLALA-----GVCQAAKLV 19
DB      148 YDVTLTMANAKDYGVAQERKRV 169

RESULT 8
US-10-467-657-7476
; Sequence 7476, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega

```

```

; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7476
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7476

Query Match      33.2%; Score 38.5; DB 6; Length 341;
Best Local Similarity 45.5%; Pred. No. 33;
Matches 10; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY      3 YDITLALA-----GVCQAAKLV 19
DB      148 YDVTLTMANAKDYGVAQERKRV 169

RESULT 9
US-10-467-657-4418
; Sequence 4418, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4418
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4418

Query Match      32.8%; Score 38; DB 6; Length 282;
Best Local Similarity 88.9%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      14 QAAKLVQOF 22
DB      151 KAAKLVQOF 159

RESULT 10
US-11-156-084-358
; Sequence 358, Application US/11156084
; Publication No. US20060010515A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; TITLE OF INVENTION: Controlled expression of cytokinin biosynthetic genes leads to
; TITLE OF INVENTION: agronomically interesting phenotypes
; FILE REFERENCE: (38-21)
; CURRENT APPLICATION NUMBER: US/11/156,084
; CURRENT FILING DATE: 2005-06-17
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 358
; LENGTH: 311
; TYPE: PRT

```

ORGANISM: Geobacter sulfurreducens PCA
US-11-156-084-358

Query Match 32.3%; Score 37.5; DB 7; Length 311;
Best Local Similarity 39.1%; Pred. No. 44;
Matches 9; Conservative 5; Mismatches 6; Indels 3; Gaps 1;

Qy 2 YVDITLALAGVC---QAAKLVQO 21
Db 249 YKEVCAYLAGTCSLDEARELWQR 271

RESULT 11
US-11-055-822-94
Sequence 94, Application US/11055822
Publication No. US20050260707A1

GENERAL INFORMATION:
APPLICANT: Pompejus, Markus

APPLICANT: Krogger, Burkhard
APPLICANT: Schröder, Hartwig

APPLICANT: Zelder, Oskar

APPLICANT: Haberhauser, Gregor

TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING

FILE REFERENCE: BGI-121CPCN

CURRENT FILING DATE: 2005-02-11

PRIOR FILING DATE: 2005-02-11

PRIOR FILING DATE: 2000-06-23

PRIOR FILING DATE: 2000-06-23

PRIOR FILING DATE: 1999-06-25

PRIOR FILING DATE: 60/142,101

PRIOR FILING DATE: 1999-07-02

PRIOR FILING DATE: 60/148,613

PRIOR FILING DATE: 1999-08-12

PRIOR FILING DATE: 60/187,970

PRIOR FILING DATE: 2000-03-09

PRIOR FILING DATE: 1999-07-01

PRIOR FILING DATE: 1999-07-01

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

APPLICANT: MONACI Elisabetha
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE: BGI-121CPCN

CURRENT FILING DATE: 2003-08-11

PRIOR FILING DATE: 2003-08-11

PRIOR FILING DATE: 2001-02-12

NUMBER OF SEQ ID NOS: 9218

SOFTWARE: Seqwin99, version 1.04

SEQ ID NO 1542

LENGTH: 54

TYPE: PRT

ORGANISM: Neisseria gonorrhoeae

US-10-467-657-1542

Query Match 31.9%; Score 37; DB 6; Length 54;
Best Local Similarity 28.6%; Pred. No. 8.4;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 3 YDITLALAGVCQAAKLVQQA 23
Db 11 FQTFVSGLCFVSGILRFA 31

RESULT 13
US-11-055-822-926
Sequence 926, Application US/11055822
Publication No. US20050260707A1

GENERAL INFORMATION:
APPLICANT: Pompejus, Markus

APPLICANT: Krogger, Burkhard

APPLICANT: Schröder, Hartwig

APPLICANT: Zelder, Oskar

APPLICANT: Haberhauser, Gregor

TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING

FILE REFERENCE: BGI-121CPCN

CURRENT FILING DATE: 2005-02-11

PRIOR FILING DATE: 2005-02-11

PRIOR FILING DATE: 2000-06-23

PRIOR FILING DATE: 2000-06-23

PRIOR FILING DATE: 1999-06-25

PRIOR FILING DATE: 60/142,101

PRIOR FILING DATE: 1999-07-02

PRIOR FILING DATE: 60/148,613

PRIOR FILING DATE: 1999-08-12

PRIOR FILING DATE: 60/187,970

PRIOR FILING DATE: 2000-03-09

PRIOR FILING DATE: 1999-07-01

PRIOR FILING DATE: 1999-07-01

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

APPLICANT: MONACI Elisabetha
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

FILE REFERENCE: BGI-121CPCN

CURRENT FILING DATE: 2003-08-11

PRIOR FILING DATE: 2003-08-11

PRIOR FILING DATE: 2001-02-12

NUMBER OF SEQ ID NOS: 9218

SOFTWARE: Seqwin99, version 1.04

SEQ ID NO 1542

LENGTH: 54

TYPE: PRT

ORGANISM: Neisseria gonorrhoeae

US-10-467-657-1542

Query Match 31.9%; Score 37; DB 6; Length 54;
Best Local Similarity 28.6%; Pred. No. 8.4;
Matches 6; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 3 YDITLALAGVCQAAKLVQQA 23
Db 11 FQTFVSGLCFVSGILRFA 31

RESULT 13
US-11-055-822-926
Sequence 926, Application US/11055822
Publication No. US20050260707A1

GENERAL INFORMATION:
APPLICANT: Pompejus, Markus

APPLICANT: Krogger, Burkhard

APPLICANT: Schröder, Hartwig

APPLICANT: Zelder, Oskar

APPLICANT: Haberhauser, Gregor

TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING

FILE REFERENCE: BGI-121CPCN

CURRENT FILING DATE: 2005-02-11

PRIOR FILING DATE: 2005-02-11

PRIOR FILING DATE: 2000-06-23

PRIOR FILING DATE: 2000-06-23

PRIOR FILING DATE: 1999-06-25

PRIOR FILING DATE: 60/142,101

PRIOR FILING DATE: 1999-07-02

PRIOR FILING DATE: 60/148,613

PRIOR FILING DATE: 1999-08-12

PRIOR FILING DATE: 60/187,970

PRIOR FILING DATE: 2000-03-09

PRIOR FILING DATE: 1999-07-01

PRIOR FILING DATE: 1999-07-01

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

PRIOR FILING DATE: 1999-07-08

Db 188 DHYDVSAATVGVVDELL 206

RESULT 14

US-11-055-822-924

/ Sequence 924, Application US/11055822
/ Publication No. US20050260707A1

/ GENERAL INFORMATION:

/ APPLICANT: Pompeius, Markus

/ APPLICANT: Kroeger, Burkhard

/ APPLICANT: Schroder, Hartwig

/ APPLICANT: Zelder, Oskar

/ APPLICANT: Haberhauser, Gregor

/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING

/ FILE REFERENCE: BGI-121CPN

/ CURRENT APPLICATION NUMBER: US/11/055,822

/ PRIOR FILING DATE: 2005-02-11

/ PRIOR APPLICATION NUMBER: 09/606,740

/ PRIOR FILING DATE: 2000-06-23

/ PRIOR APPLICATION NUMBER: 60/141,031

/ PRIOR FILING DATE: 1999-06-25

/ PRIOR APPLICATION NUMBER: 60/142,101

/ PRIOR FILING DATE: 1999-07-02

/ PRIOR APPLICATION NUMBER: 60/148,613

/ PRIOR FILING DATE: 1999-08-12

/ PRIOR APPLICATION NUMBER: 60/187,970

/ PRIOR FILING DATE: 2000-03-09

/ PRIOR APPLICATION NUMBER: DE 19930476.9

/ PRIOR FILING DATE: 1999-07-01

/ PRIOR APPLICATION NUMBER: DE 19931415.2

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931418.7

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931419.5

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

/ PRIOR APPLICATION NUMBER: DE 19931420.9

/ PRIOR FILING DATE: 1999-07-08

LENGTH: 696

TYPE: PRT

ORGANISM: Helicobacter pylori J99

US-11-052-554A-100

Query Match

Best Local Similarity 31.9%; Score 37; DB 7; Length 696;

Matches 11; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 1 NYVD-----ITLALAGVCGAATVQQA 23

Db 237 NYSGGSSIPQIKTSSVDAENLLQQA 265

Search completed: January 23, 2006, 11:53:35

Job time : 5.1 secs

US-11-055-822-924

/ ORGANISM: Corynebacterium glutamicum

/ TYPE: PRT

/ LENGTH: 390

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

/ SEQ ID NO 924

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838A-2
Perfect score: 116
Sequence: 1 NYDITLALAGVQAKLVQCPA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	100.0	23	US-10-665-184-2	Sequence 2, Appl1
2	116	100.0	23	US-10-942-300-2	Sequence 2, Appl1
3	116	100.0	23	US-10-501-838A-2	Sequence 2, Appl1
4	116	100.0	203	US-10-665-184-60	Sequence 60, Appl1
5	116	100.0	203	US-10-942-300-60	Sequence 60, Appl1
6	116	100.0	203	US-10-501-838A-54	Sequence 54, Appl1
7	103	88.8	23	US-10-665-184-3	Sequence 3, Appl1
8	103	88.8	23	US-10-942-300-3	Sequence 3, Appl1
9	103	88.8	23	US-10-501-838A-3	Sequence 3, Appl1
10	103	88.8	24	US-10-665-184-25	Sequence 25, Appl1
11	103	88.8	24	US-10-942-300-25	Sequence 25, Appl1
12	103	88.8	24	US-10-501-838A-25	Sequence 25, Appl1
13	103	88.8	25	US-10-501-838A-19	Sequence 19, Appl1
14	103	88.8	25	US-10-501-838A-23	Sequence 23, Appl1
15	103	88.8	26	US-10-501-838A-21	Sequence 21, Appl1
16	103	88.8	29	US-10-501-838A-18	Sequence 18, Appl1
17	103	88.8	30	US-10-665-184-22	Sequence 22, Appl1
18	103	88.8	30	US-10-942-300-22	Sequence 22, Appl1
19	103	88.8	30	US-10-501-838A-20	Sequence 20, Appl1
20	103	88.8	30	US-10-501-838A-22	Sequence 22, Appl1
21	103	88.8	31	US-10-665-184-35	Sequence 35, Appl1
22	103	88.8	31	US-10-942-300-35	Sequence 35, Appl1
23	103	88.8	213	US-10-501-838A-35	Sequence 35, Appl1
24	103	88.8	213	US-10-665-184-61	Sequence 61, Appl1
25	103	88.8	213	US-10-942-300-61	Sequence 61, Appl1
26	103	88.8	213	US-10-501-838A-55	Sequence 55, Appl1
27	97	83.6	23	US-10-665-184-26	Sequence 26, Appl1

28	97	83.6	23	US-10-942-300-26	Sequence 26, Appl1
29	97	83.6	23	US-10-501-838A-26	Sequence 26, Appl1
30	97	83.6	30	US-10-665-184-36	Sequence 36, Appl1
31	97	83.6	30	US-10-942-300-36	Sequence 36, Appl1
32	97	83.6	30	US-10-501-838A-36	Sequence 36, Appl1
33	92	79.3	23	US-10-665-184-1	Sequence 1, Appl1
34	92	79.3	23	US-10-942-300-1	Sequence 1, Appl1
35	92	79.3	23	US-10-501-838A-1	Sequence 1, Appl1
36	92	79.3	205	US-10-665-184-59	Sequence 59, Appl1
37	92	79.3	205	US-10-942-300-59	Sequence 59, Appl1
38	92	79.3	205	US-10-501-838A-53	Sequence 53, Appl1
39	90	77.6	22	US-10-665-184-27	Sequence 27, Appl1
40	90	77.6	22	US-10-942-300-27	Sequence 27, Appl1
41	90	77.6	22	US-10-501-838A-27	Sequence 27, Appl1
42	90	77.6	22	US-10-665-184-37	Sequence 37, Appl1
43	90	77.6	29	US-10-942-300-37	Sequence 37, Appl1
44	90	77.6	29	US-10-501-838A-37	Sequence 37, Appl1
45	89	76.7	23	US-10-665-184-29	Sequence 29, Appl1

ALIGNMENTS

RESULT 1
US-10-665-184-2
Sequence 2, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
FILE REFERENCE: 24348-501CIP
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 23
TYPE: PRT
ORGANISM: Pasteurella multocida
US-10-665-184-2

Query Match 100.0% Score 116; DB 4; Length 23;
Best Local Similarity 100.0% Pred. No. 2, 6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDITLALAGVQAKLVQCPA 23
DB 1 NYDITLALAGVQAKLVQCPA 23

RESULT 2
US-10-942-300-2
Sequence 2, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-503
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 2
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Pasteurella multocida
US-10-942-300-2

Query Match 100.0%; Score 116; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITTLAAGVCOAKLVQQFA 23
Db 1 NYDITTLAAGVCOAKLVQQFA 23

RESULT 3
US-10-501-838A-2

;; Sequence 2, Application US/10501838A
;; Publication No. US20050215478A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel A.
;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501 NATL
;; CURRENT APPLICATION NUMBER: US/10/501,838A
;; PRIOR FILING DATE: 2004-07-19
;; PRIOR APPLICATION NUMBER: PCT/IB03/00968
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: US 60/355,396
;; NUMBER OF SEQ ID NOS: 72
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 2
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Pasteurella multocida
US-10-501-838A-2

Query Match 100.0%; Score 116; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITTLAAGVCOAKLVQQFA 23
Db 1 NYDITTLAAGVCOAKLVQQFA 23

RESULT 4
US-10-665-184-60

;; Sequence 60, Application US/10665184
;; Publication No. US20040146549A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel
;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501CIP
;; CURRENT APPLICATION NUMBER: US/10/665,184
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: PCT/IB03/00968
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: 60/355,396
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 60
;; LENGTH: 203
;; TYPE: PRT
;; ORGANISM: Pasteurella multocida

US-10-665-184-60

Query Match 100.0%; Score 116; DB 4; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITTLAAGVCOAKLVQQFA 23
Db 3 NYDITTLAAGVCOAKLVQQFA 25

RESULT 5
US-10-942-300-60

;; Sequence 60, Application US/10942300
;; Publication No. US20050136103A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel
;; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-503
;; CURRENT APPLICATION NUMBER: US/10/942,300
;; PRIOR FILING DATE: 2004-09-16
;; PRIOR APPLICATION NUMBER: 10/665,184
;; PRIOR FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: 10/664,989
;; PRIOR FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: 60/503,615
;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 60
;; LENGTH: 203
;; TYPE: PRT
;; ORGANISM: Pasteurella multocida
US-10-942-300-60

Query Match 100.0%; Score 116; DB 5; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITTLAAGVCOAKLVQQFA 23
Db 3 NYDITTLAAGVCOAKLVQQFA 25

RESULT 6
US-10-501-838A-54

;; Sequence 54, Application US/10501838A
;; Publication No. US20050215478A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel A.
;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501 NATL
;; CURRENT APPLICATION NUMBER: US/10/501,838A
;; PRIOR FILING DATE: 2004-07-19
;; PRIOR APPLICATION NUMBER: PCT/IB03/00968
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: 60/355,396
;; PRIOR FILING DATE: 2002-02-07
;; NUMBER OF SEQ ID NOS: 72
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 54
;; LENGTH: 203
;; TYPE: PRT
;; ORGANISM: Pasteurella multocida
US-10-501-838A-54

Query Match 100.0%; Score 116; DB 5; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
US-10-501-838A-21
: Sequence 21, Application US/10501838A

Publication No. US20050215478A1
 GENERAL INFORMATION:
 APPLICANT: Ben-Sasson, Shmuel A.
 TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
 TITLE OF INVENTION: Biological Barrier
 FILE REFERENCE: 24348-501 NATL
 CURRENT APPLICATION NUMBER: US/10/501,838A
 CURRENT FILING DATE: 2004-07-19
 PRIOR APPLICATION NUMBER: PCT/IB03/00968
 PRIOR FILING DATE: 2003-02-07
 PRIOR APPLICATION NUMBER: US 60/355,396
 PRIOR FILING DATE: 2002-02-07
 NUMBER OF SEQ ID NOS: 72
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 21
 LENGTH: 26
 TYPE: PRT
 ORGANISM: Artificial sequence
 FEATURE:
 OTHER INFORMATION: Synthetic: penetrating peptide
 FEATURE:
 NAME/KEY: MISC_FEATURE
 LOCATION: (26)..(26)
 OTHER INFORMATION: wherein heparin is coupled to the penetrating peptide via the
 OTHER INFORMATION: free amino group of the lysine residue
 US-10-501-838A-21

Query Match 88.8%; Score 103; DB 5; Length 26;
 Best Local Similarity 82.6%; Pred. No. 3.7e-09;
 Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYVDITLALGVCOAAKLVQQFA 23
 DB 1 NYVDITLALGVCOAAKLVQQFA 23

Search completed: January 23, 2006, 11:52:32
 Job time : 61 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-3

Perfect score: 114
Sequence: 1 NYDYTLALAGICQSRRLVQQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Genesegp1980a:.*
2: genesegp1990a:.*
3: genesegp2000a:.*
4: genesegp2001a:.*
5: genesegp2002a:.*
6: genesegp2003a:.*
7: genesegp2003b:.*
8: genesegp2004a:.*
9: genesegp2005a:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	23	ADBI6888	Bescherich
2	114	100.0	23	AEBO8220	Bescherich
3	114	100.0	24	ADBI6908	Bescherich
4	114	100.0	24	AEBO8242	Bescherich
5	114	100.0	25	ADBI6922	Bescherich
6	114	100.0	25	ADBI6906	Bescherich
7	114	100.0	26	ADBI6904	E. coli pe
8	114	100.0	29	ADBI6921	E. coli pe
9	114	100.0	30	ADBI6903	Bescherich
10	114	100.0	30	ADBI6905	Bescherich
11	114	100.0	30	AEBO8239	Penetrati
12	114	100.0	31	ADBI6918	Bescherich
13	114	100.0	31	AEBO8252	Penetrati
14	114	100.0	31	AEBO8263	Bescherich
15	114	100.0	213	ABO61272	Klebsiell
16	108	94.7	217	ADBI6909	Bescherich
17	108	94.7	223	ADBI6909	Bescherich
18	108	94.7	230	AEBO8243	Bescherich
19	108	94.7	230	AEBO8253	Penetrati
20	103	90.4	23	ADBI6887	Penetrati
21	103	90.4	23	AEBO8219	Paasteurel
22	103	90.4	203	AEBO8262	Paasteurel
23	101	88.6	22	ADBI6910	Bescherich
24	101	88.6	22	AEBO8244	Bescherich

25	101	88.6	29	6	ADBI6920	Bescherich
26	101	88.6	29	9	AEBO8254	Penetrati
27	98	86.0	23	6	ADBI6912	Haemophil
28	98	86.0	23	9	AEBO8246	Haemophil
29	98	86.0	29	6	ADBI6916	Bescherich
30	98	86.0	29	6	AEBO8250	Penetrati
31	97	85.1	209	6	ABW67069	Phototrab
32	95	83.3	23	6	ADBI6886	Haemophil
33	95	83.3	23	9	AEBO8218	Haemophil
34	95	83.3	205	4	AAW50230	Haemophil
35	95	83.3	205	9	AEBO8261	Bacterial
36	88	77.2	216	7	ADFO7708	Buchnera
37	88	70.2	23	6	ADBI6890	Buchnera
38	80	70.2	23	9	AEBO8252	Buchnera
39	80	70.2	211	9	AEBO8265	Buchnera
40	72	63.2	23	6	ADBI6889	Vibrio ch
41	72	63.2	23	9	AEBO8221	Vibrio ch
42	72	63.2	204	9	AEBO8264	Vibrio ch
43	51	44.0	468	4	AAU08337	Drosophill
44	49	43.0	313	4	AAU08337	Bovine su
45	49	43.0	313	6	ABG73541	Bovine su

ALIGNMENTS

RESULT 1
ADBI6888 standard; peptide; 23 AA.
ADBI6888:
20-NOV-2003 (first entry)
Bescherichia coli YCFC penetrating peptide 3.
penetrating peptide; epithelial; endothelial; tight junction; diabetes;
infertility; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
osteopathic; cytoskeletal; nootropic.
Bescherichia coli.
WO2003066859-A2.
14-AUG-2003.
07-FEB-2003; 2003WO-IB000968.
07-FEB-2002; 2002US-035396P.
(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
Ben-Sasson SA, Cohen E;
WPI; 2003-697452/66.
New penetrating peptide, useful for preparing a composition for treating
or preventing e.g. endocrine disorders.
Claim 2; Page 14; 60pp; English.
This invention relates to a novel peptide sequences capable of
translocating across a biological barrier. Furthermore, it refers to
methods that use these peptides to facilitate penetration of a
biologically active effector molecule such as a drug or other therapeutic
agent across biological barriers e.g. epithelial or endothelial cells
sealed by tight junctions. This peptide is derived from a bacterial
toxin, an integral membrane or extracellular protein and can comprise an
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
or enzyme. The effector molecule, however, can comprise for example
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiatherosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from YCFC of
 CC Escherichia coli and is penetrating peptide 3 of the invention.

XX Sequence 23 AA;

Query Match 100.0%; Score 114; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.5e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NYDITTLALAGICQSARLVQOLA 23
 1 NYDITTLALAGICQSARLVQOLA 23

RESULT 2

ADB08220 standard; peptide; 23 AA.

AC AEB08220;

DT 25-AUG-2005 (first entry)

XX Escherichia coli YCFC penetrating peptide 3, SEQ ID NO: 3.

XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antihypertensive;
 XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; antiviral;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antineumatic; cyostatic;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.

OS Escherichia coli.

PN US2005136103-A1.

XX 23-JUN-2005.

PD 16-SEP-2004; 2004US-00942300.

XX 17-SEP-2003; 2003US-00664989.

PR 17-SEP-2003; 2003US-00665184.

PF 17-SEP-2003; 2003US-0503615P.

XX (BENS/) BEN-SASSON S A.

PA (COHEN/) COHEN E.

PI Ben-Sasson SA, Cohen E;

XX WPI; 2005-444089/45.

XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 53; SEQ ID NO 3; 59pp; English.

XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.

XX Sequence 23 AA;

Query Match 100.0%; Score 114; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.5e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NYDITTLALAGICQSARLVQOLA 23
 1 NYDITTLALAGICQSARLVQOLA 23

RESULT 3

ADB16908 standard; peptide; 24 AA.

AD 16-SEP-2004;

DT 20-NOV-2003 (first entry)

XX Escherichia coli YCFC penetrating peptide 30.

XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 XX infertility; hormone; vitamin deficiency; neurodegenerative;
 XX cardiovascular; haematological; endocrine disorder; obesity;
 XX neoplastic disease; neuroprotective; cardiac; antihypertensive;
 XX osteopathic; cyostatic; neurotropic.

OS Escherichia coli.

PN WO2003066859-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003WO-IB000968.

XX 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PA Ben-Sasson SA, Cohen E;

XX WPI; 2003-697452/66.

XX New penetrating peptide, useful for preparing a composition for treating

PT or preventing e.g. endocrine disorders.
 XX
 XX Claim 2, Page 15; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC antibiotic, antidiabetic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiatherosclerotic, osteopathic,
 CC cytostatic or neurotropic activities. This peptide is from YCFC of
 CC *Escherichia coli* and is penetrating peptide 30 of the invention.
 XX
 SQ Sequence 24 AA:
 XX
 Query Match 100.0%; Score 114; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.6e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 NYDITLALAGICQSARLVQOLA 23
 Db 2 NYDITLALAGICQSARLVQOLA 24
 XX
 RESULT 4
 AEB08242 ID AEB08242 standard; peptide; 24 AA.
 XX
 AC AEB08242;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DE *Escherichia coli* YCFC penetrating peptide 30, SEQ ID NO: 25.
 XX
 KW Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW demerit; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotoxic;
 KW genitourinary disease; haematological disease; anaemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 KW
 OS *Escherichia coli*.
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.

XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 XX (COHR/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 XX
 DR WPI; 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 Claim 46; SEQ ID NO 25; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, renal failure,
 CC disorders, obesity, vitamin deficiencies, renal disorders, immunologic and
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the *Escherichia coli* YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 SQ Sequence 24 AA:
 XX
 Query Match 100.0%; Score 114; DB 9; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.6e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 NYDITLALAGICQSARLVQOLA 23
 Db 2 NYDITLALAGICQSARLVQOLA 24
 XX
 RESULT 5
 ADB16922 ID ADB16922 standard; peptide; 25 AA.
 XX
 AC ADB16922;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE *Escherichia coli* peptide 3 linked to recombinant human insulin.
 XX
 KW recombinant human insulin; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cytostatic; neurotropic; penetrating peptide; diabetes.
 XX
 OS Synthetic.
 XX
 XX *Escherichia coli*.

```

FH Key location/Qualifiers
FT Peptide 1..23 "Penetrating peptide 3"
FT Modified-site 25
FT /note= "Recombinant human insulin peptide coupled to the C-terminus"
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX
XX WPI; 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Example 8; Page 42; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells
XX sealed by tight junctions. This peptide is derived from a bacterial
XX toxin, an integral membrane or extracellular protein and can comprise an
XX anticonagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX or enzyme. The effector molecule, however, can comprise for example
XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
XX factors. The penetrating peptide is useful for the treatment of various
XX conditions including diabetes, infertility, hormone and vitamin
XX deficiencies, neurodegenerative, cardiovascular, haematological and
XX endocrine disorders, as well as obesity and neoplastic disease.
XX Accordingly, the peptides of this invention can be used in compositions
XX that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
XX cyrostatic or nootropic activities. This peptide sequence consists of the
XX Escherichia coli penetrating peptide 3 coupled to recombinant human
XX insulin in the absence of a detachable linker peptide as a cleavage site,
XX used to deliver insulin across mucosal epithelia as a treatment for
XX diabetes, in an exemplification of the invention.
XX
XX Sequence 25 AA;
XX
XX Query Match 100.0%; Score 114; DB 6; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-11;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 NYDITTLAAGICOSARLVQOLA 23
XX 1 NYDITTLAAGICOSARLVQOLA 23
XX
XX RESULT 6
XX ADB16906
XX ID ADB16906 standard; peptide; 25 AA.
XX
XX ADB16906;
XX
XX 20-NOV-2003 (first entry)
XX
XX Escherichia coli peptide 3 linked to linearised insulin receptor peptide.
XX linearised insulin receptor; epithelial; endothelial; tight junction;
XX diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
XX cardiovascular; haematological; endocrine disorder; obesity;

```

```

KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
KW osteopathic; cyrostatic; nootropic; penetrating peptide; diabetes.
XX
XX Synthetic.
XX Escherichia coli.
XX
XX Key location/Qualifiers
XX Peptide 1..23 "Penetrating peptide 3"
XX Modified-site 25
XX /note= "Linearised insulin receptor peptide coupled to the C-terminus"
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX
XX WPI; 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Example 9; Page 43; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells
XX sealed by tight junctions. This peptide is derived from a bacterial
XX toxin, an integral membrane or extracellular protein and can comprise an
XX anticonagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX or enzyme. The effector molecule, however, can comprise for example
XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
XX factors. The penetrating peptide is useful for the treatment of various
XX conditions including diabetes, infertility, hormone and vitamin
XX deficiencies, neurodegenerative, cardiovascular, haematological and
XX endocrine disorders, as well as obesity and neoplastic disease.
XX Accordingly, the peptides of this invention can be used in compositions
XX that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
XX cyrostatic or nootropic activities. This peptide sequence consists of the
XX Escherichia coli penetrating peptide 3 coupled to a linearised insulin
XX receptor, used to deliver insulin across mucosal epithelia as a treatment
XX for diabetes, in an exemplification of the invention.
XX
XX Sequence 25 AA;
XX
XX Query Match 100.0%; Score 114; DB 6; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-11;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 NYDITTLAAGICOSARLVQOLA 23
XX 1 NYDITTLAAGICOSARLVQOLA 23
XX
XX RESULT 7
XX ADB16904
XX ID ADB16904 standard; peptide; 26 AA.
XX
XX ADB16904;
XX
XX 20-NOV-2003 (first entry)
XX

```

DE E_coli penetrating peptide 3 linked to heparin without a cleavage site.
 XX heparin; epithelial; endothelial; tight junction; diabetes; infertility;
 KW hormone; vitamin deficiency; neurodegenerative; cardiovascular;
 KW haematological; endocrine disorder; obesity; neoplastic disease;
 KW neuroprotective; cardiant; antiarteriosclerotic; osteopathic; cyostatic;
 KW neotropic; penetrating peptide.
 XX Synthetic.
 OS Escherichia coli.
 XX
 XX Key Location/Qualifiers
 XX Peptide 1..23
 XX /note= "Penetrating peptide 3"
 XX Modified-site 26
 XX /note= "Heparin peptide coupled to the C-terminus"
 XX
 XX WO2003066859-A2.
 XX
 XX 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX
 XX Example 10; Page 44; 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dargargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, haematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX Accordingly, the peptides of this invention can be used in compositions
 XX that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 XX cyostatic or neotropic activities. This peptide sequence consists of the
 XX Escherichia coli penetrating peptide 3 coupled to heparin in the absence
 XX of a detachable linker peptide as a cleavage site, used to deliver low
 XX molecular weight heparin across mucosal epithelia in an exemplification
 XX of the invention.
 XX
 XX Sequence 26 AA;
 XX
 XX Query Match 100.0%; Score 114; DB 6; Length 26;
 XX Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 NYDITLALAGICQSARLVQOLA 23
 XX ||||||||||||||||||||
 XX 1 NYDITLALAGICQSARLVQOLA 23
 XX
 XX RESULT 8
 XX ADB16921
 XX ID ADB16921 standard; peptide; 29 AA.

XX ADB16921;
 XX
 XX 20-NOV-2003 (first entry)
 XX
 XX E_coli peptide 3 linked to recombinant human insulin via cleavage site.
 XX
 XX recombinant human insulin; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
 KW osteopathic; cyostatic; neotropic; detachable linker;
 KW penetrating peptide; diabetes.
 XX Synthetic.
 OS Escherichia coli.
 XX
 XX Key Location/Qualifiers
 XX Peptide 1..23
 XX /note= "Penetrating peptide 3"
 XX Cleavage-site 26..29
 XX /note= "Detachable linker peptide"
 XX Modified-site 29
 XX /note= "Recombinant human insulin peptide coupled to the
 XX C-terminus"
 XX
 XX WO2003066859-A2.
 XX
 XX 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX
 XX Example 8; Page 42; 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dargargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, haematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX Accordingly, the peptides of this invention can be used in compositions
 XX that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 XX cyostatic or neotropic activities. This peptide sequence consists of the
 XX Escherichia coli penetrating peptide 3 coupled to recombinant human
 XX insulin via the detachable linker peptide as a cleavage site, used to
 XX deliver insulin across mucosal epithelia as a treatment for diabetes, in
 XX an exemplification of the invention.
 XX
 XX Sequence 29 AA;
 XX
 XX Query Match 100.0%; Score 114; DB 6; Length 29;
 XX Best Local Similarity 100.0%; Pred. No. 1.9e-11;
 XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX

QY 1 NYDITLALAGICOSARLVQOLA 23
 DB 1 NYDITLALAGICOSARLVQOLA 23

RESULT 9
 ADB16903
 ID ADB16903 standard; peptide; 30 AA.
 AC ADB16903;
 XX 20-NOV-2003 (first entry)
 DT
 DE E_col1 penetrating peptide 3 linked to heparin through a cleavage site.
 XX
 KW heparin; epithelial; endothelial; tight junction; diabetes; infertility; hormone; vitamin deficiency; neurodegenerative; cardiovascular; haematological; endocrine disorder; obesity; neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic; osteopathic; cytostatic; nootropic; detachable linker; penetrating peptide.
 KW
 OS Synthetic.
 OS Escherichia coli.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..23
 FT /note= "Penetrating peptide 3"
 FT 26..29
 FT Cleavage-site /note= "Detachable linker peptide"
 FT 30
 FT Modified-site /note= "Heparin peptide coupled to the C-terminus"
 FT
 FT
 XX WO2003066859-A2.
 XX 14-AUG-2003.
 XX 07-FEB-2003; 2003WO-IB000968.
 XX 07-FEB-2002; 2002US-0355396P.
 PR
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA Ben-Sasson SA, Cohen E;
 PI WPI, 2003-697452/66.
 DR
 XX New penetrating peptide, useful for preparing a composition for treating or preventing e.g. endocrine disorders.
 PT
 PT
 XX Example 10; Page 44; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, haematological and endocrine disorders, as well as obesity and neoplastic disease. Accordingly, the peptides of this invention can be used in compositions that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic, cytostatic or nootropic activities. This peptide sequence consists of the Escherichia coli penetrating peptide 3 coupled to heparin via the detachable linker peptide as a cleavage site, used to deliver low molecular weight heparin across mucosal epithelia in an exemplification

CC of the invention.
 XX
 SQ Sequence 30 AA;
 Query Match 100.0%; Score 114; DB 6; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDITLALAGICOSARLVQOLA 23
 DB 1 NYDITLALAGICOSARLVQOLA 23

RESULT 10
 ADB16905
 ID ADB16905 standard; peptide; 30 AA.
 AC ADB16905;
 XX 20-NOV-2003 (first entry)
 DT
 DE Escherichia coli peptide 3 coupled to imaging compound linker, IBM-002.
 XX
 KW epithelial; endothelial; tight junction; diabetes; infertility; hormone; vitamin deficiency; neurodegenerative; cardiovascular; haematological; endocrine disorder; obesity; neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic; osteopathic; cytostatic; nootropic; imaging linker; penetrating peptide; IBM-002.
 KW
 OS Synthetic.
 OS Escherichia coli.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT 30
 FT Modified-site /note= "C-terminal amide"
 FT
 FT
 XX WO2003066859-A2.
 XX 14-AUG-2003.
 XX 07-FEB-2003; 2003WO-IB000968.
 XX 07-FEB-2002; 2002US-0355396P.
 PR
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA Ben-Sasson SA, Cohen E;
 PI WPI, 2003-697452/66.
 DR
 XX New penetrating peptide, useful for preparing a composition for treating or preventing e.g. endocrine disorders.
 PT
 PT
 XX Example 1; Page 33; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, haematological and endocrine disorders, as well as obesity and neoplastic disease. Accordingly, the peptides of this invention can be used in compositions

CC that have neuroprotective, cardiact, antiarteriosclerotic, osteopahic,
CC cytoabatic or nootropic activities. This peptide sequence is IBN-002,
CC which consists of the Escherichia coli penetrating peptide 3 coupled
CC the imaging linker peptide used in an exemplification of the invention.

SQ Sequence 30 AA;

Query Match	100.0%	Score 114	DB 6	Length 30
Similarity	100.0%	Prod. No.	2e-11	
Best Local				
Matches 23	Conservative 0	Mismatches 0	Gaps 0	

[illegible]

RESULT 11
AEB08239
ID AEB08239 standard; peptide; 30 AA

AC	ABB08239;
XX	
DT	25-AUG-2005 (first entry)

XX Penetrating peptide SEQ: 22 used in composition for mucosal vaccination
 XX
 XX Pharmaceutical, therapeutic, endocrine disease, endocrine-gen., diabetes
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurologic; neurological disease; parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antiatherosclerotic;
 XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX genitourinary disease; hematological disease; antineum; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; virocidic;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antineumatic; cytostatic;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.

OS Unidentified

FN	key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

PT	/note= "N-terminal acylated"
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

PT /note= "Optionally C-terminal amide, optionally the free
PT amino group of lysine is acylated with a fatty acid"

PN US2005136103-A1.

PD 23-JUN-2005

PR 16-SEP-2004; 2004US-00942300.

PR 17-SEP-2003; 2003US-00664989

PR 17-SEP-2003; 2003US-0503615P

PA (BENS/) BEN-SASSON S. A.

PA (COHE/) COHEN E.

PA (COHE/) COHEN E.

PI Ben-Sabbon SA, Cohen E;

DR WPI, 2005-444089/45

X

PT Composition used for translocating effectors across barrier such as
PT epithelial cells during treatment of e.g. endocrine disorders comprises
PT effector sequentially coupled with counter ion and hydrophobic agent.
XX
PS Claim 63; SEQ ID NO 22; 59pp; English.

Claim 63; SEQ ID NO 22; 59pp; English.

The present invention relates to a pharmaceutical composition of penetrating peptides for transepithelial delivery of effector. The invention comprises the effector sequentially coupled with a counter ion and at least one hydrophobic agent, where the effector is selectively encapsulated into a complex. The invention is useful for translocating effectors across a biological barrier such as epithelial cells and endothelial cells during treatment and prevention of disease or pathological conditions (including endocrine disorders, diabetes, infertility, hormone deficiencies, osteoporosis, ophthalmological disorders, neurodegenerative disorders, Alzheimer's disease, dementia, Parkinson's disease, multiple sclerosis, Huntington's disease, cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypocoagulable states, coronary disease, cerebrovascular events, metabolic disorders, obesity, vitamin deficiencies, renal disorders, renal failure, hematological disorders, anemia of different entities, immunologic and rheumatologic disorders, autoimmune diseases, immune deficiencies, infectious diseases, viral infections, bacterial infections, fungal infections, parasitic infections, neoplastic diseases, multi-factorial disorders, impotence, chronic pain, depression, different fibrosis states and short stature) and for mucosal vaccination against anthrax and hepatitis B. The present sequence is a penetrating peptide (1BW-002) used in the composition for mucosal vaccination using a counter ion and a penetrating peptide.

SQ Sequence 30 AA;

Query Match	100.0%	Score 114	DB 9	Length 30
Best Local Similarity	100.0%	Pred. No. 2e-11		
Matches 23	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Oy 1 NYDITLALAGICQSARLVQOLA 23
 |||||
 Db 1 NYDITLALAGICQSARLVQOLA 23

RESULT 12

ID ADB16918 standard; peptide; 31 AA.

AC ADB16918;

DE Escherichia coli peptide 3 coupled to imaging compound linker, IBW-002V1

KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
vitamin deficiency; neurodegenerative; cardiovascular; haematological;

KW cardiant; antiart

XX

OS Escherichia coli

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

/Note= "N-terminal acetyl"

PT 311

PN WO2003066859

PD 14-AUG-2003.

PF 07-FEB-2003; 2003WO-IB0000968

PR

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2003-697452/66.
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 XX Example 3; Page 40; 60pp; English.
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antidiabetic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antidiabetic, osteopathic,
 CC cyostatic or neurotrophic activities. This peptide sequence is IBW-002VI,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 XX
 XX Sequence 31 AA;
 SQ

Query Match 100.0%; Score 114; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYDYITLALAGICOSARLVQOLA 23
 |||||
 DB 2 NYDYITLALAGICOSARLVQOLA 24

RESULT 13
 AEB08252 ID AEB08252 standard; peptide; 31 AA.
 XX AEB08252;
 XX 25-AUG-2005 (first entry)
 XX Penetrating peptide SEQ. 35 used in composition for mucosal vaccination.
 XX
 XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX Antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotrophic; neurological disease; Parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; vinnicide;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antineumatic; cyostatic;

KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX Unidentified.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT Misc-difference 1 /note= "N-terminal acylated"
 FT /note= "Optionally C-terminal amide, optionally the free
 FT amino group of lysine is acylated with a fatty acid"
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 PF
 XX 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.
 DR
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 63; SEQ ID NO 35; 59pp; English.
 PS

The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies, fungal
 CC infectious diseases, viral infections, bacterial infections, multi-factorial
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is a penetrating peptide (IBW-002VI)
 CC used in the composition for mucosal vaccination using a counter anion and
 CC a penetrating peptide.

Sequence 31 AA;
 Query Match 100.0%; Score 114; DB 9; Length 31;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYDYITLALAGICOSARLVQOLA 23
 |||||
 DB 2 NYDYITLALAGICOSARLVQOLA 24

RESULT 14
 AEB08263 ID AEB08263 standard; protein; 213 AA.
 XX

AC AEB08263;
 XX 25-AUG-2005 (first entry)
 DT XX
 XX Escherichia coli protein, SEQ ID NO: 61.
 DE XX
 XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer disease;
 XX neurotropic; neurological disease; parkinsons disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntington chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; virologic;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antineumatic; cytostatic;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX Escherichia coli.
 OS
 XX US2005136103-A1.
 PN
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 PP
 XX 17-SEP-2003; 2003US-00664989.
 PR
 XX 17-SEP-2003; 2003US-00665184.
 PR
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.
 DR
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Disclosure; SEQ ID NO 61; 59pp; English.
 PS
 XX The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli protein

CC containing penetrating peptide at N-terminal end.
 XX
 XX SQ Sequence 213 AA;

Query Match 100.0%; Score 114; DB 9; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.9e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSAKLVQOLA 23
 |||||
 Db 4 NYDITLALAGICQSAKLVQOLA 26

RESULT 15
 ABO61272
 ID ABO61272 standard; protein; 217 AA.
 XX
 XX AC ABO61272;
 XX
 XX DT 29-JUL-2004 (first entry)
 XX
 XX DE Klebsiella pneumoniae polypeptide seqid 7789.
 XX
 XX KW Recombinant expression vector; transcription regulatory element;
 XX KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 XX OS Klebsiella pneumoniae.
 XX
 XX PN US610836-B1.
 XX
 XX PD 26-AUG-2003.
 XX
 XX PF 27-JAN-2000; 2000US-00489039.
 XX
 XX PR 29-JAN-1999; 99US-0117747P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX
 XX PA Breton GL, Osborne M;
 XX
 XX PI WPI; 2003-895346/82.
 XX
 XX DR N-PSDB; ACH94823.
 XX
 XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 XX PS Disclosure; SEQ ID NO 7789; 932pp; English.
 XX
 XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 XX SQ Sequence 217 AA;

Query Match 96.5%; Score 110; DB 7; Length 217;
 Best Local Similarity 91.3%; Pred. No. 9e-10;
 Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSAKLVQOLA 23
 |||||
 Db 8 NYDITLALAGICQSAKLVQOLA 30

Search completed: January 23, 2006, 09:20:47
 Job time : 70.2 secs

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-3

Sequence: 1 NYVDITLALAGICQSARLVQOLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	114	100.0	2 AH0199	conserved hypothe
2	114	100.0	2 S19211	ycfc protein - Esc
3	114	100.0	2 D90829	hypothetical prote
4	114	100.0	2 B85687	hypothetical prote
5	111	97.4	2 AB0647	conserved hypothe
6	95	83.3	2 164155	hypothetical prote
7	80	70.2	2 D84960	hypothetical prote
8	72	63.2	2 D82237	conserved hypothe
9	48	42.1	2 157957	neurokinin 2 recep
10	48	42.1	2 S20303	neurokinin 2 recep
11	48	42.1	2 A36737	neurokinin 2 recep
12	47	41.2	2 B91207	heat shock protein
13	47	41.2	2 G65170	heat shock protein
14	47	41.2	2 D86053	heat shock protein
15	47	41.2	2 384	neurokinin 2 recep
16	47	41.2	2 1500516	neurokinin 2 recep
17	47	41.2	2 Q01059	neurokinin 2 recep
18	46	40.4	2 156595	neurokinin 2 recep
19	46	40.4	2 AF0961	heat shock protein
20	46	40.4	2 AC0496	heat shock protein
21	46	40.4	2 A44156	lipopolysaccharide
22	46	40.4	2 S70671	lipopolysaccharide
23	45	39.5	2 JCI1065	coat protein - bee
24	45	39.5	2 AH0891	adenyl-transferase
25	44	38.6	2 B82680	conserved hypothe
26	44	38.6	2 A10326	cel operon repres
27	43	37.7	2 B71543	hypothetical prote
28	43	37.7	2 B81698	conserved hypothe
29	43	37.7	2 T00567	hypothetical prote
30	43	37.7	2 B96653	hypothetical prote

30	43	37.7	759	2	D70422	cellulose synthase
31	43	37.7	847	2	G95843	conserved hypothe
32	42	36.8	98	2	T22503	hypothetical prote
33	42	36.8	273	2	AC3277	transposase BME102
34	42	36.8	336	2	G71954	lipopolysaccharide
35	42	36.8	340	2	G64554	lipopolysaccharide
36	42	36.8	433	2	T25946	hypothetical prote
37	41.5	36.4	510	2	D96741	hypothetical prote
38	41.5	36.4	1215	2	T43916	chitinase A (impor
39	41	36.0	260	2	A34178	bacteriorhodopsin
40	41	36.0	274	2	S76154	hypothetical prote
41	41	36.0	317	2	AF0973	lipopolysaccharide
42	41	36.0	319	2	A40619	lipopolysaccharide
43	41	36.0	330	2	D86038	hepoxyl transfera
44	41	36.0	330	2	C91191	hepoxyl transfera
45	41	36.0	429	2	T24922	hypothetical prote

ALIGNMENTS

RESULT 1
AH0199
Conserved hypothetical protein YP01637 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence revision 02-Nov-2001 #ext_change 09-Jul-2004
C/Accession: AH0199
R/Parikh, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
Hill, M.; Rutherford, S.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527 (2000)
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB00047; MIMD:21470437; PMID:11586360
A/Accession: AH0199
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1208 <KIR>
A/Cross-references: UNIPROT:Q82F06; UNIPARC:UP100000CD828; GB:AL590842; PIDN:CAC90459.1
C/Genetics:
A/Gene: YP01637

Query Match 100.0%; Score 114; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYVDITLALAGICQSARLVQOLA 23
DB 4 NYVDITLALAGICQSARLVQOLA 26

RESULT 2
S19211
ycfc protein - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #ext_change 09-Jul-2004
C/Accession: S19211; A64858
R/Green, S.M.; Drablie, W.T.
Submitted to the EMBL Data Library, May 1991
A/Description: Molecular analysis of the purB-phoP region of Escherichia coli K12.
A/Reference number: S19210
A/Accession: S19211
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-213 <GRE>
A/Cross-references: UNIPROT:P25746; UNIPARC:UP1000013A673; EMBL:X59307; NID:G42562; PID
A/Experiment: Source: strain K-12
R/Battcher, F.R.; Plunkett III, G.; Bloch, C.A.; Berna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Sings, Y.
Science 277, 1453-1462 (1997)
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MIMD:97426617; PMID:9278503
A/Accession: A64858
A/Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-213 <BLAT>
A:Cross-references: UNIPARC:UPI000013A673; GB:AE000213; GB:U00096; NID:g1787371; PIDN:AA
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ynfC
A:Start codon: GTG

Query Match 100.0%; Score 114; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NYDITTLAAGICGASARLVQOLA 23
DB 4 NYDITTLAAGICGASARLVQOLA 26

RESULT 3
D90829
hypothetical protein EC61604 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D90829
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A95629; MUID:21156231; PMID:11258796
A:Status: preliminary
A:Accession: D90829
A:Molecule type: DNA
A:Residues: 1-213 <HAY>
A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:BA000007; PIDN:BA035027.1;
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC61604

Query Match 100.0%; Score 114; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NYDITTLAAGICGASARLVQOLA 23
DB 4 NYDITTLAAGICGASARLVQOLA 26

RESULT 4
B85687
hypothetical protein ynfC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85687
R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:AE005174; NID:g12514786; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ynfC

Query Match 100.0%; Score 114; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 NYDITTLAAGICGASARLVQOLA 23
DB 4 NYDITTLAAGICGASARLVQOLA 26

RESULT 5
AB0647
conserved hypothetical protein STY1273 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typh
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0647
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0647
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-215 <PAR>
A:Cross-references: UNIPARC:UPI0000059F77; GB:AL513382; PIDN:CAD08357.1; PID:g16502402;
C:Genetics:
A:Gene: STY1273

Query Match 97.4%; Score 111; DB 2; Length 215;
Best Local Similarity 95.7%; Pred. No. 1.2e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 NYDITTLAAGICGASARLVQOLA 23
DB 6 NYDITTLAAGICGASARLVQOLA 28

RESULT 6
164155
hypothetical protein H10638 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: 164155
R:Flaischman, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Feldman, J
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fritchman, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: 164155
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-205 <TTGR>
A:Cross-references: UNIPROT:P44796; UNIPARC:UPI000013A674; GB:U32747; GB:LA2023; NID:g15
A>Note: best homolog was a hypothetical protein from Escherichia coli

Query Match 83.3%; Score 95; DB 2; Length 205;
Best Local Similarity 78.3%; Pred. No. 4.8e-08;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 NYDITTLAAGICGASARLVQOLA 23
DB 3 NYDITTLAAGICGASARLVQOLA 25

RESULT 7
D84960
hypothetical protein [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: D84960
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A>Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: D84960

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: EC64626

C:Superfamily: alpha-crystallin-related small heat shock protein

Query Match

Best Local Similarity 41.2%; Score 47; DB 2; Length 144;

Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSARLVQ 20

DB 45 NHYRITLALAGFROEDLEIQ 64

RESULT 13

G65170

heat shock protein lbpb - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C:Accession: G65170; B45245

R:Blattner, P.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: G65170

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-144 <BLAT>

A:Cross-references: UNIPROT:Q8XC04; UNIPARC:UPI00000D052C; GB:AE000445; GB:U00096; NID:9

R:Allen, S.P.; Polazzi, J.O.; Gierse, J.K.; Easton, A.M.

J. Bacteriol. 174, 6938-6947, 1992

A:Title: Two novel heat shock genes encoding proteins produced in response to heterolog

A:Reference number: A45245; MUID:93015757; PMID:1356969

A:Accession: B45245

A:Status: preliminary

A:Molecule type: DNA; protein

A:Residues: 3-144 <ALL>

A:Cross-references: UNIPARC:UPI000012D125; GB:M94104; NID:g147368; PIDN:AA24425.1; PID:

A:Note: Sequence extracted from NCBI backbone (NCBI:116883, NCBI:P:116883)

C:Genetics:

A:Gene: hsls

C:Superfamily: alpha-crystallin-related small heat shock protein

Query Match

Best Local Similarity 41.2%; Score 47; DB 2; Length 144;

Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSARLVQ 20

DB 45 NHYRITLALAGFROEDLEIQ 64

RESULT 14

D86053

heat shock protein [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: D86053

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaeser, J.D.; Rose, D.J.; Mayhew

iller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D86053

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-144 <STO>

A:Cross-references: UNIPROT:Q8XC04; UNIPARC:UPI00000D052C; GB:AE005174; NID:g12518527; F

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: lbpb

C:Superfamily: alpha-crystallin-related small heat shock protein

Query Match

Best Local Similarity 41.2%; Score 47; DB 2; Length 144;

Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSARLVQ 20

DB 45 NHYRITLALAGFROEDLEIQ 64

RESULT 15

S00516

neurokinin 2 receptor - bovine

N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)

C:Species: Bos primigenius taurus (cattle)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: S00516

R:Maui, Y.; Nakayama, K.; Tamaki, H.; Harada, Y.; Kuno, M.; Nakashima, S.

Nature 329, 836-838, 1997

A:Title: cDNA cloning of bovine substance-K receptor through oocyte expression system.

A:Reference number: S00516; MUID:88039072; PMID:2823146

A:Accession: S00516

A:Molecule type: mRNA

A:Residues: 1-384 <MAS>

A:Cross-references: UNIPROT:P05363; UNIPARC:UPI00001301BE; EMBL:X06295; NID:g746; PIDN:C

C:Superfamily: neurokinin 1 receptor

C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane pr

F:33-59/Domain: transmembrane #status predicted <TM1>

F:70-93/Domain: transmembrane #status predicted <TM2>

F:111-159/Domain: transmembrane #status predicted <TM3>

F:150-169/Domain: transmembrane #status predicted <TM4>

F:193-222/Domain: transmembrane #status predicted <TM5>

F:252-275/Domain: transmembrane #status predicted <TM6>

F:288-310/Domain: transmembrane #status predicted <TM7>

F:11,19/Binding site: carbohydrate (Aen) (covalent) #status predicted

F:106-181/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 41.2%; Score 47; DB 1; Length 384;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSA 16

DB 69 NYFIVNLALDLQMAA 84

Search completed: January 23, 2006, 09:32:48
Job time : 9.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 Seconds
(without alignments)
114.551 Million cell updates/sec

Title: US-10-501-838A-3
Perfect score: 114
Sequence: 1 NYDITLALGICQASRLVQQLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents, AA: *
1: /cgn2_6/prodata/1/1aa/5.COMB.pep: *
2: /cgn2_6/prodata/1/1aa/6.COMB.pep: *
3: /cgn2_6/prodata/1/1aa/H.COMB.pep: *
4: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep: *
5: /cgn2_6/prodata/1/1aa/RE.COMB.pep: *
6: /cgn2_6/prodata/1/1aa/backfillset.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	96.5	217	2	US-09-489-039A-7789 Sequence 7789, Ap
2	88	77.2	216	2	US-09-543-681A-7993 Sequence 7993, Ap
3	51	44.7	138	2	US-09-270-767-58510, A Sequence 58510, A
4	51	44.7	138	2	US-09-270-767-43172 Sequence 43172, A
5	48	42.1	369	2	US-07-937-609-19 Sequence 19, Appl
6	48	42.1	369	2	US-08-029-170-19 Sequence 19, Appl
7	48	42.1	369	2	US-09-443-745-19 Sequence 19, Appl
8	47	41.2	384	1	US-08-103-170-10 Sequence 10, Appl
9	47	41.2	387	1	US-08-196-989B-14 Sequence 14, Appl
10	47	41.2	387	2	US-08-760-936-14 Sequence 14, Appl
11	47	41.2	387	2	US-09-225-024-14 Sequence 14, Appl
12	47	41.2	398	2	US-09-489-039A-10685 Sequence 527, App
13	46	40.4	231	2	US-09-489-039A-10685 Sequence 10685, A
14	46	40.4	231	2	US-09-830-433A-32 Sequence 32, Appl
15	44	38.6	462	1	US-08-471-033-20 Sequence 20, Appl
16	44	38.6	462	1	US-08-471-044-20 Sequence 20, Appl
17	44	38.6	462	1	US-08-463-483A-20 Sequence 20, Appl
18	44	38.6	462	1	US-08-471-046A-20 Sequence 20, Appl
19	44	38.6	462	1	US-08-470-566B-20 Sequence 20, Appl
20	44	38.6	462	1	US-08-469-334-20 Sequence 20, Appl
21	44	38.6	462	1	US-09-300-529-20 Sequence 20, Appl
22	42.5	37.3	427	2	US-09-248-796A-17203 Sequence 17203, A
23	42.5	37.3	762	2	US-09-228-986-114 Sequence 114, App
24	42.5	37.3	762	2	US-10-101-464A-114 Sequence 114, App
25	41.5	36.8	823	2	US-09-252-991A-23655 Sequence 23655, A
26	41.5	36.8	717	2	US-09-107-532A-6350 Sequence 6350, Ap
27	41	36.0	60	2	US-09-248-796A-27407 Sequence 27407, A

28	41	36.0	114	2	US-09-902-540-10640 Sequence 10640, A
29	41	36.0	231	2	US-09-830-433A-71 Sequence 71, Appl
30	41	36.0	309	2	US-09-902-540-12828 Sequence 12828, A
31	41	36.0	400	2	US-09-543-681A-4587 Sequence 4587, Ap
32	41	36.0	400	2	US-09-328-352-7817 Sequence 7817, Ap
33	41	36.0	439	2	US-09-902-540-14780 Sequence 14780, A
34	41	36.0	588	2	US-09-605-703B-280 Sequence 280, App
35	41	36.0	588	2	US-09-605-703B-282 Sequence 282, App
36	41	36.0	741	2	US-09-252-991A-27062 Sequence 27062, A
37	40	35.1	191	2	US-09-489-039A-10634 Sequence 10634, A
38	40	35.1	303	2	US-09-543-681A-7924 Sequence 7924, Ap
39	40	35.1	311	2	US-09-489-039A-12640 Sequence 12640, A
40	40	35.1	324	2	US-09-489-039A-10801 Sequence 10801, A
41	40	35.1	398	2	US-09-543-681A-7046 Sequence 7046, Ap
42	40	35.1	409	1	US-08-360-672-6 Sequence 6, Appl
43	40	35.1	409	5	US-09-640-305-6 Sequence 6, Appl
44	40	35.1	410	1	US-08-088-633-2 Sequence 2, Appl
45	40	35.1	410	1	US-08-245-756-2 Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-489-039A-7789
Sequence 7789, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489, 039A
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117, 747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7789
LENGTH: 217
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match          96.5%; Score 110; DB 2; Length 217;
Best Local Similarity 91.3%; Pred. No. 1.8e-10;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

CY      1 NYDITLALGICQASRLVQQLA 23
      |||||:|||||:|||||
Db      8 NYDITLALGVQCARLVQQLA 30

RESULT 2
US-09-543-681A-7993
Sequence 7993, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543, 681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128, 706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7993
LENGTH: 216
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7993

Query Match          77.2%; Score 88; DB 2; Length 216;
```

Best Local Similarity 73.9%; Pred. No. 8-6e-07;
Matches 17; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQASRLVQOLA 23
Db 11 DFRDITLALAGICQASRLVQOLA 33

RESULT 3

US-09-270-767-58510
; Sequence 58510, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 58510
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-58510

Query Match 44.7%; Score 51; DB 2; Length 138;
Best Local Similarity 58.8%; Pred. No. 0.8;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQASR 17
Db 24 NYADIEIALAGIADKAR 40

RESULT 4

US-09-270-767-43172
; Sequence 43172, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 43172
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43172

Query Match 44.7%; Score 51; DB 2; Length 433;
Best Local Similarity 58.8%; Pred. No. 3;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQASR 17
Db 319 NYADIEIALAGIADKAR 335

RESULT 5

US-07-937-609-19
; Sequence 19, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; SOFTWARE: Patentin Release #1.0, Version #1.25
; NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIND
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: rat substance K receptor
US-07-937-609-19

Query Match 42.1%; Score 48; DB 1; Length 369;
Best Local Similarity 56.2%; Pred. No. 7.9;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSA 16
Db 69 NYFIINLALADLCMAA 84

RESULT 6

US-08-029-170-19
; Sequence 19, Application US/08029170
; Patent No. 6169173
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIH
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-08-029-170-19

Query Match 42.1%; Score 48; DB 2; Length 369;
Best Local Similarity 56.2%; Pred. No. 7.9;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQA 16
DB 69 NYFIITLALADLCMAA 84

RESULT 7
US-09-443-745-19
Sequence 19, Application US/09443745
Patent No. 6706493
GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033

FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIH
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-09-443-745-19

Query Match 42.1%; Score 48; DB 2; Length 369;
Best Local Similarity 56.2%; Pred. No. 7.9;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQA 16
DB 69 NYFIITLALADLCMAA 84

RESULT 8
US-08-103-170-10
Sequence 10, Application US/08103170
Patent No. 5885824
GENERAL INFORMATION:
APPLICANT: Yamada, Tadataka
TITLE OF INVENTION: Recombinant Genomic Clones Encoding
TITLE OF INVENTION: Histamine H1, H2, and H3 Receptors, Methods For Production
TITLE OF INVENTION: Thereof, and Proteins Encoded Therefrom
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCLELLAND, WATERS & NEUSTADT,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/633,060
FILING DATE: 24-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-017-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 24885 OPAT UR
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Bos taurus
US-08-103-170-10

Query Match 41.2%; Score 47; DB 1; Length 384;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSA 16
||:|||||:|:
Db 69 NYFIWNLADLCMAA 84

RESULT 9
US-08-196-989B-14
Sequence 14, Application US/08196989B
Patent No. 5585476
GENERAL INFORMATION:
APPLICANT: MacLennan, A. John
TITLE OF INVENTION: Molecular Cloning and Expression of
TITLE OF INVENTION: G-Protein Coupled Receptors
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/196,989B
FILING DATE: 15-FEB-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Lloyd, Jeff
REGISTRATION NUMBER: 35,589
REFERENCE/DOCKET NUMBER: MAC-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-196-989B-14

Query Match 41.2%; Score 47; DB 1; Length 387;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSA 16
||:|||||:|:
Db 69 NYFIWNLADLCMAA 84

RESULT 10
US-08-760-936-14
Sequence 14, Application US/08760936
Patent No. 5856443
GENERAL INFORMATION:

APPLICANT: MacLennan, A. John
TITLE OF INVENTION: Molecular Cloning and Expression of
TITLE OF INVENTION: G-Protein Coupled Receptors
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,936
FILING DATE: December 6, 1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: MAC-100C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-760-936-14

Query Match 41.2%; Score 47; DB 1; Length 387;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSA 16
||:|||||:|:
Db 69 NYFIWNLADLCMAA 84

RESULT 11
US-09-225-024-14
Sequence 14, Application US/09225024
Patent No. 6518414
GENERAL INFORMATION:
APPLICANT: MacLennan, A. John
TITLE OF INVENTION: Molecular Cloning and Expression of
TITLE OF INVENTION: G-Protein Coupled Receptors
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/225,024
FILING DATE: 04-JAN-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/760,936
FILING DATE: 6-DEC-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/196,989
FILING DATE: 15-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: MAC-100C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-225-024-14

Query Match 41.2%; Score 47; DB 2; Length 387;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSA 16
||:|||||:
DB 69 NYPIVLAADLCMAA 84

RESULT 12
US-09-826-509-527
Sequence 527, Application US/09826509
Patent No. 6806054
GENERAL INFORMATION:
APPLICANT: Lehmann-Bruhnsma, Karin
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: Patentin Version 2.1
SEQ ID NO 527
LENGTH: 398
TYPE: PRT
ORGANISM: Homo sapiens
US-09-826-509-527

Query Match 41.2%; Score 47; DB 2; Length 398;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSA 16
||:|||||:
DB 69 NYPIVLAADLCMAA 84

RESULT 13
US-09-489-039A-10685
Sequence 10685, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10685
LENGTH: 169
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10685

Query Match 40.4%; Score 46; DB 2; Length 169;
Best Local Similarity 71.4%; Pred. No. 7;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQ 14
||:|||||:
DB 70 NHRITLALAGFRQ 83

RESULT 14
US-09-830-433A-32
Sequence 32, Application US/09830433A
Patent No. 6835384
GENERAL INFORMATION:
APPLICANT: AJJAME et al.
TITLE OF INVENTION: Nucleic acids and polypeptides specific for pathogenic
FILE REFERENCE: P07180US00/BAS
CURRENT APPLICATION NUMBER: US/09/830,433A
CURRENT FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: FR 98 13 693
PRIOR FILING DATE: 1998-10-30
NUMBER OF SEQ ID NOS: 129
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 32
LENGTH: 231
TYPE: PRT
ORGANISM: Neisseria meningitidis
US-09-830-433A-32

Query Match 40.4%; Score 46; DB 2; Length 231;
Best Local Similarity 56.2%; Pred. No. 10;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 YYDITLALAGICQSA 17
||:|||||:
DB 195 YYDLAPAPICRQAR 210

RESULT 15
US-08-471-033-20
Sequence 20, Application US/08471033
Patent No. 5770696
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziele, Michael G
APPLICANT: Mullins, Marcha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalin M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Betrich, Juan J
TITLE OF INVENTION: No. 5770696el Peestical Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.308
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,033
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-471-033-20

```

```

Query Match      38.6%; Score 44; DB 1; Length 462;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY      1 NYDITLALAGICQ 14 .
Db      107 NYKETTSMAGSCE 120

```

Search completed: January 23, 2006, 09:35:44
 Job time : 16.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 11:32:56 ; Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838a-3
Sequence: 1 NYDDTLALAGICQARLVQQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: /pubseq/1/seqs/AA_New.*
2: /pubseq/2/seqs/US08_NEW_PUB.pep.*
3: /pubseq/3/seqs/US06_NEW_PUB.pep.*
4: /pubseq/4/seqs/US07_NEW_PUB.pep.*
5: /pubseq/5/seqs/US09_NEW_PUB.pep.*
6: /pubseq/6/seqs/US10_NEW_PUB.pep.*
7: /pubseq/7/seqs/US11_NEW_PUB.pep.*
8: /pubseq/8/seqs/US12_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	37.7	42	US-10-512-295A-3	Sequence 3, Appl1
2	43	37.7	185	US-10-453-372-896	Sequence 896, App
3	43	37.7	194	US-10-453-372-894	Sequence 894, App
4	43	37.7	218	US-10-453-372-892	Sequence 892, App
5	41	36.0	305	US-10-520-820-1	Sequence 1, Appl1
6	40.5	35.5	365	US-11-000-463-713	Sequence 241, App
7	40.5	35.5	365	US-11-000-463-713	Sequence 713, App
8	40	35.1	63	US-10-467-657-6638	Sequence 6638, App
9	39	34.2	1015	US-10-957-569-51	Sequence 51, Appl1
10	38.5	33.8	447	US-10-858-730-219	Sequence 219, App
11	38	33.3	203	US-10-453-372-890	Sequence 890, App
12	38	33.3	655	US-10-055-877-40	Sequence 40, Appl1
13	38	33.3	678	US-10-055-877-202	Sequence 202, App
14	38	33.3	687	US-10-055-877-199	Sequence 199, App
15	38	33.3	687	US-10-055-877-201	Sequence 201, App
16	37	32.5	357	US-10-821-234-1569	Sequence 1569, App
17	36.5	32.0	447	US-11-055-822-94	Sequence 94, Appl1
18	36.5	32.0	447	US-10-467-657-4660	Sequence 4660, App
19	36	31.6	228	US-10-467-657-1242	Sequence 1242, App
20	36	31.6	301	US-10-467-657-7356	Sequence 7356, App
21	36	31.6	350	US-10-467-657-3256	Sequence 3256, App
22	36	31.6	434	US-10-467-657-6894	Sequence 6894, App
23	36	31.6	473	US-11-069-642-5	Sequence 5, Appl1
24	36	31.6	491	US-10-763-712A-4	Sequence 4, Appl1
25	36	31.6	919	US-10-858-730-206	Sequence 206, App

26	36	31.6	1268	7	US-11-052-554A-1	Sequence 1, Appl1
27	35.5	31.1	341	6	US-10-467-657-154	Sequence 154, App
28	35.5	31.1	341	6	US-10-467-657-5822	Sequence 5822, App
29	35.5	31.1	341	6	US-10-467-657-7476	Sequence 7476, App
30	35.5	31.1	884	6	US-10-878-556A-9	Sequence 9, Appl1
31	35.5	31.1	974	6	US-10-821-234-1152	Sequence 1152, App
32	35.5	31.1	1259	6	US-10-467-657-5510	Sequence 5510, App
33	35	30.7	250	6	US-10-793-626-2310	Sequence 2310, App
34	35	30.7	270	6	US-10-467-657-7100	Sequence 7100, App
35	35	30.7	293	6	US-11-067-671-15	Sequence 15, Appl1
36	35	30.7	305	7	US-11-000-463-391	Sequence 391, App
37	35	30.7	305	7	US-11-000-463-863	Sequence 863, App
38	35	30.7	332	6	US-10-454-437-308	Sequence 308, App
39	35	30.7	356	6	US-10-793-626-2054	Sequence 2054, App
40	35	30.7	445	6	US-11-194-246-345	Sequence 345, App
41	35	30.7	502	6	US-10-063-703-158	Sequence 158, App
42	35	30.7	502	7	US-11-102-240-158	Sequence 158, App
43	35	30.7	651	7	US-11-124-368A-184	Sequence 184, App
44	35	30.7	1588	7	US-11-052-554A-280	Sequence 280, App
45	34.5	30.3	522	7	US-11-055-822-1100	Sequence 1100, App

ALIGNMENTS

RESULT 1
US-10-512-295A-3
; Sequence 3, Application US/10512295A
; Publication No. US20050245727A1
GENERAL INFORMATION:
; APPLICANT: Baltzer, Lars
; APPLICANT: Dolphin, Gunnar
; APPLICANT: Liedberg, Bo
; APPLICANT: Lundstrom, Ingemar
; TITLE OF INVENTION: NOVEL POLYPEPTIDE SCAFFOLDS AND USE THEREOF
; FILE REFERENCE: 5848.181USWO
; CURRENT APPLICATION NUMBER: US/10/512,295A
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/SE03/00507
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: SE 0200968-6
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Synthetic Peptide KES3
US-10-512-295A-3
Query Match 37.7%; Score 43; DB 6; Length 42;
Best Local Similarity 52.9%; Pred. No. 0.6;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Db 7 LALAGICQARLVQQLA 23
16 LAARPCDAQLAQLA 32
RESULT 2
US-10-453-372-896
; Sequence 896, Application US/10453372
; Publication No. US20060003323A1
GENERAL INFORMATION:
; APPLICANT: Alsbjork, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23

```
;; PRIOR APPLICATION NUMBER: 60/185967
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 09/823187
;; PRIOR FILING DATE: 2001-03-29
;; PRIOR APPLICATION NUMBER: 60/195792
;; PRIOR FILING DATE: 2000-03-10
;; PRIOR APPLICATION NUMBER: 09/839446
;; PRIOR FILING DATE: 2001-03-19
;; PRIOR APPLICATION NUMBER: 60/199476
;; PRIOR FILING DATE: 2000-03-25
;; PRIOR APPLICATION NUMBER: 09/863776
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: 60/208263
;; PRIOR FILING DATE: 2000-05-31
;; PRIOR APPLICATION NUMBER: 09/939398
;; PRIOR FILING DATE: 2001-08-24
;; PRIOR APPLICATION NUMBER: 60/227800
;; PRIOR FILING DATE: 2000-08-25
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1609
;; SOFTWARE: CurSeqList version 0.1
;; SEQ ID NO: 896
;; LENGTH: 185
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-453-372-896
```

```
Query Match      37.7%; Score 43; DB 6; Length 185;
Best Local Similarity 52.4%; Pred. No. 2.8;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
Oy      3 YDITLALAGICQSARLVQOLA 23
Db      67 YDSLALDGHIGSARALMVVA 87
```

RESULT 3

```
US-10-453-372-894
; Sequence 894, Application US/10453372
; Publication No. US2006000323A1
```

GENERAL INFORMATION:

APPLICANT: Alsebrook, et al.

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

FILE REFERENCE: 21402-589 A

CURRENT APPLICATION NUMBER: US/10/453,372

PRIOR FILING DATE: 2003-06-03

PRIOR APPLICATION NUMBER: 09/789390

PRIOR FILING DATE: 2001-02-23

PRIOR APPLICATION NUMBER: 60/185967

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 09/823187

PRIOR FILING DATE: 2001-03-29

PRIOR APPLICATION NUMBER: 60/195792

PRIOR FILING DATE: 2000-03-10

PRIOR APPLICATION NUMBER: 09/839446

PRIOR FILING DATE: 2001-03-19

PRIOR APPLICATION NUMBER: 60/199476

PRIOR FILING DATE: 2000-03-25

PRIOR APPLICATION NUMBER: 09/863776

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: 60/208263

PRIOR FILING DATE: 2000-05-31

PRIOR APPLICATION NUMBER: 09/939398

PRIOR FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: 60/227800

PRIOR FILING DATE: 2000-08-25

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1609

SOFTWARE: CurSeqList version 0.1

SEQ ID NO: 894

LENGTH: 194

TYPE: PRT

ORGANISM: Homo sapiens

US-10-453-372-894

```
Query Match      37.7%; Score 43; DB 6; Length 194;
Best Local Similarity 52.4%; Pred. No. 3;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
Oy      3 YDITLALAGICQSARLVQOLA 23
Db      47 YDSLALDGHIGSARALMVVA 67
```

RESULT 4

```
US-10-453-372-892
; Sequence 892, Application US/10453372
; Publication No. US2006000323A1
```

GENERAL INFORMATION:

APPLICANT: Alsebrook, et al.

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

FILE REFERENCE: 21402-589 A

CURRENT APPLICATION NUMBER: US/10/453,372

PRIOR FILING DATE: 2003-06-03

PRIOR APPLICATION NUMBER: 09/789390

PRIOR FILING DATE: 2001-02-23

PRIOR APPLICATION NUMBER: 60/185967

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 09/823187

PRIOR FILING DATE: 2001-03-29

PRIOR APPLICATION NUMBER: 60/195792

PRIOR FILING DATE: 2000-03-10

PRIOR APPLICATION NUMBER: 09/839446

PRIOR FILING DATE: 2001-03-19

PRIOR APPLICATION NUMBER: 60/199476

PRIOR FILING DATE: 2000-03-25

PRIOR APPLICATION NUMBER: 09/863776

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: 60/208263

PRIOR FILING DATE: 2000-05-31

PRIOR APPLICATION NUMBER: 09/939398

PRIOR FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: 60/227800

PRIOR FILING DATE: 2000-08-25

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1609

SOFTWARE: CurSeqList version 0.1

SEQ ID NO: 892

LENGTH: 218

TYPE: PRT

ORGANISM: Homo sapiens

US-10-453-372-892

```
Query Match      37.7%; Score 43; DB 6; Length 218;
Best Local Similarity 52.4%; Pred. No. 3.4;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
Oy      3 YDITLALAGICQSARLVQOLA 23
Db      71 YDSLALDGHIGSARALMVVA 91
```

RESULT 5

```
US-10-520-820-1
; Sequence 1, Application US/10520820
; Publication No. US2006000393A1
```

GENERAL INFORMATION:

APPLICANT: MUTHARILIS S.A.

TITLE OF INVENTION: Pathogenicity determinants which can be used as targets for devel.

TITLE OF INVENTION: means for preventing and controlling bacterial infections and/or

FILE REFERENCE: 1621

CURRENT APPLICATION NUMBER: US/10/520,820

PRIOR FILING DATE: 2005-01-07

PRIOR APPLICATION NUMBER: PCT/EP2003/008209

PRIOR FILING DATE: 2003-07-09

PRIOR APPLICATION NUMBER: FR 0208636
PRIOR FILING DATE: 2002-07-09
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 305
TYPE: PRT
ORGANISM: Escherichia coli
US-10-520-820-1

Query Match 36.0%; Score 41; DB 6; Length 305;
Best Local Similarity 42.9%; Pred. No. 10;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 YDITLALAGICGASRLVQOLA 23
DB 69 YDAVIDAQGLVKSALVTRLA 89

RESULT 6

US-11-000-463-241
Sequence 241, Application US/11000463
Publication No. US20050266423A1
GENERAL INFORMATION:
APPLICANT: Tang, Y Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Chen, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIP4CN
CURRENT APPLICATION NUMBER: US/11/000,463
PRIOR FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 241
LENGTH: 365
TYPE: PRT
ORGANISM: Homo sapiens
US-11-000-463-241

Query Match 35.5%; Score 40.5; DB 7; Length 365;
Best Local Similarity 43.5%; Pred. No. 15;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 4 DITLALAGIC---QSARLVQOLA 23
DB 67 DLMGLAGFCGSMASGHLFKOMA 89

RESULT 7
US-11-000-463-713
Sequence 713, Application US/11000463

Publication No. US20050266423A1
GENERAL INFORMATION:
APPLICANT: Tang, Y Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Chen, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIP4CN
CURRENT APPLICATION NUMBER: US/11/000,463
PRIOR FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 713
LENGTH: 365
TYPE: PRT
ORGANISM: Homo sapiens
US-11-000-463-713

Query Match 35.5%; Score 40.5; DB 7; Length 365;
Best Local Similarity 43.5%; Pred. No. 15;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 4 DITLALAGIC---QSARLVQOLA 23
DB 67 DLMGLAGFCGSMASGHLFKOMA 89

RESULT 8
US-10-467-657-6638
Sequence 6638, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SpA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
PRIOR FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 6638
LENGTH: 63
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6638

Query Match 35.1%; Score 40; DB 6; Length 63;

Best Local Similarity 53.3%; Pred. No. 2.9;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 YDITLALAGICQSA 16
Db 46 YNITVASTVKCQTA 60

RESULT 9

US-10-957-569-51
; Sequence 51, Application US/10957569
; Publication No. US20050246785A1
; GENERAL INFORMATION:
; APPLICANT: COOK, Zhihong et al.
; TITLE OF INVENTION: PROMOTER, PROMOTER CONTROL ELEMENTS, AND COMBINATIONS, AND USES
; FILE REFERENCE: 2750-1577PUS3
; CURRENT FILING DATE: 2004-09-30
; PRIOR FILING DATE: 2004-09-23
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-957-569-51

Query Match 34.2%; Score 39; DB 6; Length 1015;
Best Local Similarity 35.0%; Pred. No. 80;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 4 DITLALAGICQSA 23
Db 172 DMTLMILGCAFLVSLVIGIA 191

RESULT 10

US-10-858-730-219
; Sequence 219, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE REFERENCE: 14184-030001
; CURRENT FILING DATE: 2004-06-01
; PRIOR FILING DATE: 2004-05-30
; PRIOR FILING DATE: 2003-05-30
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 219
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-858-730-219

Query Match 33.8%; Score 38.5; DB 6; Length 447;
Best Local Similarity 33.3%; Pred. No. 41;
Matches 10; Conservative 5; Mismatches 4; Indels 11; Gaps 1;

Qy 1 NYDITL-----ALAGICQSA 19
Db 9 NYDMLKRNAGEPPEFHQVAEVLSEKLV 38

RESULT 11

US-10-453-372-890
; Sequence 890, Application US/10453372
; Publication No. US20060003223A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT FILING DATE: 2003-06-03
; PRIOR FILING DATE: 2001-02-23
; PRIOR FILING DATE: 2001-02-23
; PRIOR FILING DATE: 2000-03-01
; PRIOR FILING DATE: 2001-03-29
; PRIOR FILING DATE: 2000-03-10
; PRIOR FILING DATE: 2000-03-19
; PRIOR FILING DATE: 2001-03-19
; PRIOR FILING DATE: 2000-03-25
; PRIOR FILING DATE: 2000-03-25
; PRIOR FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-05-31
; PRIOR FILING DATE: 2001-08-24
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Curaseqdist version 0.1
; SEQ ID NO 890
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-890

Query Match 33.3%; Score 38; DB 6; Length 203;
Best Local Similarity 47.6%; Pred. No. 21;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 3 YDITLALAGICQSA 23
Db 67 YDSLLADGRPQARALMVVA 87

RESULT 12

US-10-055-877-40
; Sequence 40, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tcherev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Szytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kerkuta, Rameah
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zernhuesen, Bryan
; APPLICANT: Andrew, David


```

APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eissen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 655
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-877-40

Query Match          33.3%; Score 38; DB 6; Length 655;
Best Local Similarity 64.3%; Pred. No. 74;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      9 LAGICQSARLVQQL 22
DB      558 LVGIQVRAQLVQAL 571

RESULT 13
US-10-055-877-202
Sequence 202, Application US/10055877
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Retelli, Luca
APPLICANT: Kokuda, Ramesh
```

```

APPLICANT: Guo, Xiaojia
APPLICANT: Zernusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eissen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 202
LENGTH: 678
TYPE: PRT
ORGANISM: Oryzctolagus cuniculus
US-10-055-877-202

Query Match          33.3%; Score 38; DB 6; Length 678;
Best Local Similarity 64.3%; Pred. No. 76;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      9 LAGICQSARLVQQL 22
DB      590 LVGIQVRAQLVQAL 603

RESULT 14
US-10-055-877-199
Sequence 199, Application US/10055877
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
```

```

; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Zerhusen, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Elesen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Verneet, Corine
; APPLICANT: Tauplier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US/10/055,877
; PRIOR FILING DATE: 2002-01-19
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 199
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-877-199

Query Match          33.3%; Score 38; DB 6; Length 687;
Best Local Similarity 64.3%; Pred. No. 78;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy          9 LAGICQSAKLVQQL 22
Db          590 LVGIQRAQLVQAL 603

RESULT 15
US-10-055-877-201
; Sequence 201, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchenev, Velizar
; APPLICANT: Zhong, Mei
```

```

; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Zerhusen, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Elesen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Verneet, Corine
; APPLICANT: Tauplier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US/10/055,877
; PRIOR FILING DATE: 2002-01-19
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 201
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-055-877-201

Query Match          33.3%; Score 38; DB 6; Length 687;
Best Local Similarity 64.3%; Pred. No. 78;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy          9 LAGICQSAKLVQQL 22
Db          590 LVGIQRAQLVQAL 603
```

Search completed: January 23, 2006, 11:53:35
Job time : 5.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 seconds

(without alignments)
267,334 Million cell updates/sec

Title: US-10-501-838a-3

Perfect score: 114
Sequence: 1 NYXDITLALAGICQGARLVQQLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	208	1 Y1637 YERPE	O82f66 Yersinia pe
2	114	100.0	208	1 O66903 YERPS	O66903 Yersinia pe
3	114	100.0	213	1 YCFC_ECOL57	O8f1b7 escherichia
4	114	100.0	213	1 YCFC_ECOL6	P25766 escherichia
5	114	100.0	213	1 YCFC_ECOL1	O831f8 shigella fl
6	114	100.0	213	1 YCFC_SHITL	O827f0 salmonella
7	111	97.4	213	1 YCFC_SALTY	O57qcl salmonella
8	111	97.4	213	1 YCFC_SALTY	O57qcl salmonella
9	111	97.4	213	1 O5PMJ3 SALPA	O57qcl salmonella
10	107	93.9	213	2 O57QCL_SALCH	O57qcl salmonella
11	106	93.0	204	2 O65VAV_MANSN	O57qcl salmonella
12	105	92.1	212	2 O6D4E8 ERWCT	O57qcl salmonella
13	103	90.4	203	1 Y1850_PASWU	O57qcl salmonella
14	97	85.1	208	1 Y2805_PHOHL	O57qcl salmonella
15	95	83.3	205	2 Y638_HAETN	O57qcl salmonella
16	95	83.3	205	2 O4OM69_HAETB	O57qcl salmonella
17	95	83.3	205	2 Y1650_HAEDU	O57qcl salmonella
18	82	71.9	215	1 Y262_BUCAI	O57qcl salmonella
19	80	70.2	211	1 Y243_BUCBP	O57qcl salmonella
20	74	64.9	217	1 Y127_VIBCB	O57qcl salmonella
21	72	63.2	205	1 Y1129_VIBBP	O57qcl salmonella
22	72	63.2	205	1 Y1129_VIBBP	O57qcl salmonella
23	70	61.4	205	1 Y2927_VIBVU	O57qcl salmonella
24	69	60.5	205	2 O5E3W6_VIBF1	O57qcl salmonella
25	62	54.4	205	2 O6LTI9_PHOPR	O57qcl salmonella
26	51	44.7	468	1 O9VGQ1_DROME	O57qcl salmonella
27	50	43.9	150	1 IBPB_BRWCT	O57qcl salmonella
28	49	43.0	391	2 O5GYN3_XANOR	O57qcl salmonella
29	48	42.1	197	2 O9B8R3_RHILC	O57qcl salmonella
30	48	42.1	207	2 O60CA8_METCA	O57qcl salmonella
31	48	42.1	216	2 O6C8B5_YARLI	O57qcl salmonella

32	48	42.1	266	2 O7S4H4_NEUCR	O7S4H4 neurospora
33	48	42.1	274	2 O5CSB5_CRYPV	O5CSB5 cryptospori
34	48	42.1	274	2 O5CF88_CRYHO	O5CF88 cryptospori
35	48	42.1	372	2 O7PSM6_ANOGA	O7PSM6 anopheles g
36	48	42.1	384	1 NK2R_MESAV	P51144 mesocricetu
37	48	42.1	384	1 NK2R_MOUSE	P30549 mus musculu
38	48	42.1	384	2 O5DUB0_MERUN	O5DUB0 meriones un
39	48	42.1	384	2 O8BZV9_MOUSE	O8BZV9 mus musculu
40	48	42.1	390	1 NK2R_RAT	P16610 rattus norv
41	48	42.1	599	2 O726U4_DESVH	O726U4 desulfovibr
42	48	42.1	3223	2 O5K5U8_ECOLI	O5K5U8 escherichia
43	48	42.1	3223	2 O8L1I9_ECOLI	O8L1I9 escherichia
44	48	42.1	3223	2 O8VNR6_ECOLI	O8VNR6 escherichia
45	48	42.1	3223	2 O8VQR2_ECOLI	O8VQR2 escherichia

ALIGNMENTS

RESULT 1
Y1637 YERPE STANDARD; PRT; 208 AA.
AC O82F66; O8D0Q1;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein YP01637/Y1798/YP1767.
GN OrderedLocustNames=YP01637, Y1798, YP1767;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1] Y1637 YERPE
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CO-92 / Bivovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Tishall R.W., Holden M.T.G.,
BA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyatton P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,
RT Genome sequence of Yersinia pestis, the causative agent of plague.
RL Nature 413:523-527 (2007)
RN [2] Y1637 YERPE
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KIMS / Bivovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
DOI=10.1128/JB.184.16.4601-4611.2002;
RA Dang W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.B., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.,
RT Genome sequence of Yersinia pestis KIM.
RL J. Bacteriol. 184:4601-4611 (2002).
RN [3] Y1637 YERPE
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Bivovar Mediaevalis;
RX PubMed=15368893;
DOI=10.1128/JB.184.16.4601-4611.2002;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.,
RT Complete genome sequence of Yersinia pestis strain 91001, an isolate
of avirulent to humans.
RL DNA Res. 11:179-197 (2004).
CC -1- SIMILARITY: Belongs to the YP0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

```
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ414149; CAC90459.1; -; Genomic_DNA.
CC EMBL; AE013782; AAM85366.1; ALT_INIT; Genomic_DNA.
CC EMBL; AE017133; AAS61994.1; -; Genomic_DNA.
CC PIR; AH0199; AH0199.
CC SMR; Q82F06; 2-207.
CC HAMAP; MF_006951; -; 1.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC CONFIRMED 1 M -> MIV (in Ref. 2).
CC SEQUENCE 208 AA; 22730 MM; D0CC2B3BDB499723 CRC64;

Query Match 100.0%; Score 114; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 4,7e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NYDITTLALGICOSARLVQOLA 23
4 NYDITTLALGICOSARLVQOLA 26

RESULT 2
066903 YERP8 PRELIMINARY; PRT; 208 AA.
ID Q66903
AC Q66903
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=YERP82431;
OS Yersinia pseudotuberculosis.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Yersinia.
CC NCBI_TaxID=633;
CC [1]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=IP32953 / Serotype I;
CC PubMed=15358858; DOI=10.1073/pnas.0404012101;
CC Chui P.S.G., Camfield E., Larimer F.W., Lamerdin J., Stoutland P.O.,
CC Regala W.M., Georgescu A.M., Verges L.M., Land M.L., Motin V.L.,
CC Brubaker R.R., Fowler J., Hinebuch J., Marceau M., Medigue C.,
CC Simonet M., Chenel-Francoise V., Souza B., Dacheux D., Elliott J.M.,
CC Derbise A., Hauser L.J., Garcia E.,
CC "Insights into the evolution of Yersinia pestis through whole-genome
CC comparison with Yersinia pseudotuberculosis."
CC Proc. Natl. Acad. Sci. U.S.A. 101.13826-13831(2004).
CC EMBL; BX936398; CAH21669.1; -; Genomic_DNA.
CC SMR; Q66903; 2-207.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC KW SEQUENCE 208 AA; 22730 MM; D0CC2B3BDB499723 CRC64;

Query Match 100.0%; Score 114; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 4,7e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NYDITTLALGICOSARLVQOLA 23
4 NYDITTLALGICOSARLVQOLA 26

RESULT 3
YCF3_ECOS7 STANDARD; PRT; 213 AA.
ID YCF3_ECOS7
AC Q8X736;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycf3.

GN Name=ycf3; OrderedLocustNames=21861, Ec01604;
OS Escherichia coli O157:H7.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Escherichia.
CC NCBI_TaxID=83334;
CC [1]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=O157:H7 / BDL933 / ATCC 700927 / EHEC;
CC MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
CC Perna N.T., Plunkett G., Illi, Burland V., Mau B., Glasner J.D.,
CC Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
CC Poefal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
CC Grobeck E.J., Davis N.W., Lam A., Dimantanta B.T., Potamouzis K.,
CC Apodaca J., Ananthraman T.S., Lin J., Yen G., Schwartz D.C.,
CC Welch R.A., Blattner F.R.;
CC "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
CC Nature 409:529-533(2001).
CC [2]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
CC MEDLINE=21156231; PubMed=11258796;
CC Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
CC Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
CC Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
CC Kuhara S., Shiba T., Hattori M., Shinagawa H.;
CC "Complete genome sequence of enterohaemorrhagic Escherichia coli
CC O157:H7 and genomic comparison with a laboratory strain K-12."
CC J. Bacteriol. 183:2033-2042(2001).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE005174; AAC55958.1; -; Genomic DNA.
CC EMBL; BA000007; BAB35027.1; -; Genomic_DNA.
CC PIR; B85687; B85687.
CC PIR; B90829; D90829.
CC SMR; Q8X736; 2-213.
CC HAMAP; MF_006951; -; 1.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC KW SEQUENCE 213 AA; 22947 MM; E26EP9678C3944E2 CRC64;

Query Match 100.0%; Score 114; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 4,8e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NYDITTLALGICOSARLVQOLA 23
4 NYDITTLALGICOSARLVQOLA 26

RESULT 4
YCF3_ECOS6 STANDARD; PRT; 213 AA.
ID YCF3_ECOS6
AC Q8P1B7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycf3.
GN Name=ycf3; OrderedLocustNames=c1511;
OS Escherichia coli O6.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Escherichia.
CC NCBI_TaxID=217992;
CC [1]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
CC MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
```

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Raabso D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic *Escherichia coli*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, AE016759; AAN79980.1; ALT_INT; Genomic_DNA.
 CC SMR, Q8FIB7; 2-213.
 DR HAMAP, MF_00695; -; 1.
 DR InterPro, IPR007451; DUF489.
 DR Pfam, PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 213 AA; 22934 MW; 71C8F96F8EA37B CRC64;
 SO
 Query Match 100.0%; Score 114; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 4, 8e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYDITLALAGICQSARLVQQLA 23
 DB 4 NYDITLALAGICQSARLVQQLA 26
 ID YCFC_ECOLI STANDARD; PRT; 213 AA.
 AC P25746;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Hypothetical UPF0274 protein yycF (ORF-23).
 GN Name=ycfC; OrderedLocustNames=bl132;
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OC NCBI_TaxID=562;
 OK
 RN 11
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K12;
 RC Green S.M., Drabble W.T.;
 RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K12;
 RC MEDLINE=92104952; PubMed=1729205;
 RA He B., Smith J.M., Zaikin H.;
 RT "*Escherichia coli* purB gene: cloning, nucleotide sequence, and
 RT regulation by purR";
 RL J. Bacteriol. 174:130-136(1992).
 RL [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12 / MG1655;
 RC MEDLINE=9746617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RL [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12;
 RC MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horinchi T.;
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:1137-115(1996).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to
 CC framehifts.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, X59307; CAA11995.1; -; Genomic DNA.
 CC EMBL, W74924; -; NOT_ANNOTATED_CDS; Genomic DNA.
 CC EMBL, U00096; AAC74216.1; -; Genomic DNA.
 CC EMBL, D90748; BAA35954.1; -; Genomic DNA.
 CC EMBL, D90749; BAA35963.1; -; Genomic DNA.
 CC PIR, S19211; S19211.
 DR PDB, 1Q24; X-ray; A=2-213.
 DR PDB, 1SD1; X-ray; A=2-213.
 DR Echobase; B11321; -.
 DR EcoGene; EG11345; yycF.
 DR HAMAP, MF_00695; -; 1.
 DR InterPro, IPR007451; DUF489.
 DR Pfam, PF04356; DUF489; 1.
 KW 3D-structure; Complete proteome; Hypothetical protein.
 SQ SEQUENCE 213 AA; 22948 MW; E26EF9698C3CB42 CRC64;
 SO
 Query Match 100.0%; Score 114; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 4, 8e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYDITLALAGICQSARLVQQLA 23
 DB 4 NYDITLALAGICQSARLVQQLA 26
 ID YCFC_SHIFL STANDARD; PRT; 213 AA.
 AC O81F8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Hypothetical UPF0274 protein yycF.
 GN Name=ycfC; OrderedLocustNames=SF1151, S1234;
 OS *Shigella flexneri*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Shigella*.
 OC NCBI_TaxID=623;
 OK
 RN 11
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=301 / Serotype 2a;
 RC MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
 RT through comparison with genomes of *Escherichia coli* K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RL [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RC MEDLINE=22590274; PubMed=12704152;
 RX DOI=10.1128/JAI.71.5.2775-2786.2003;

RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Dalling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of *Shigella*
 RT flexneri serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AB005674; AA042768.1; ALT_INIT; Genomic_DNA.
 CC EMBL; AB016982; AAP1657.1; -; Genomic_DNA.
 CC SMR; Q83LF8; 2-213.
 CC DR HAMAP; MF_00695; -; 1.
 CC DR InterPro; IPR007451; DUF489.
 CC DR Pfam; PF04356; DUF489; 1.
 CC KM Complete proteome; Hypothetical protein.
 CC SEQUENCE 213 AA; 22894 MW; 0B0BEC698C29FB4D CRC64;
 SQ
 Query Match 100.0%; Score 114; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 4,8e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 NYDITLALAGICQSARLVQOLA 23
 |||||
 Db 4 NYDITLALAGICQSARLVQOLA 26
 |||||
 RESULT 7
 YCFC_SALTY STANDARD; PRT; 213 AA.
 ID YCFC_SALTY
 AC Q827H0;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC.
 GN Name=ycfC; OrderedLocNames=STY1273, t1687;
 OS Salmonella typhi.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Salmonella.
 CC NCBI_TaxID=601;
 CC
 CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 CC STRAIN=CT18;
 CC MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 CC Parikh J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 CC Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatia M.,
 CC Baker S., Bauman D., Brooks K., Chillingworth T., Connor P.,
 CC Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 CC Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 CC Krogh A., Larsen T.S., Leach S., Moule S., O'Gaora P., Parry C.,
 CC Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 CC Whitehead S., Barrett B.G.;
 CC "Complete genome sequence of a multiple drug resistant *Salmonella*
 CC enterica serovar Typhi CT18.";
 CC Nature 413:848-852(2001).
 CC [2]
 CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 CC STRAIN=Ty2 / ATCC 700931;
 CC MEDLINE=22531367; PubMed=12644504;
 CC DOI=10.1128/DB.185.7.2330-2337.2003;
 CC Deng W., Liu S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 CC Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 CC "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2
 CC and CT18.";
 CC J. Bacteriol. 185:2330-2337(2003).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AL607269; CAD08357.1; ALT_INIT; Genomic_DNA.
 CC EMBL; AB016839; AAO69312.1; ALT_INIT; Genomic_DNA.
 CC SMR; Q827H0; 2-213.
 CC DR HAMAP; MF_00695; -; 1.
 CC DR InterPro; IPR007451; DUF489.
 CC DR Pfam; PF04356; DUF489; 1.
 CC KM Complete proteome; Hypothetical protein.
 CC SEQUENCE 213 AA; 22942 MW; 63192D1C806ED58 CRC64;
 SQ
 Query Match 97.4%; Score 111; DB 1; Length 213;
 Best Local Similarity 95.7%; Pred. No. 1.5e-09;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 NYDITLALAGICQSARLVQOLA 23
 |||||
 Db 4 NYDITLALAGICQSARLVQOLA 26
 |||||
 RESULT 8
 YCFC_SALTY STANDARD; PRT; 213 AA.
 ID YCFC_SALTY
 AC Q82PZ5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC.
 GN Name=ycfC; OrderedLocNames=STY1233;
 OS Salmonella typhimurium.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Salmonella.
 CC NCBI_TaxID=602;
 CC
 CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 CC STRAIN=LT2 / SGC1412 / ATCC 700720;
 CC MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
 CC McDaniel M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 CC Courtney L., Portolillo S., Ali J., Dante M., Du P., Hou S., Layman D.,
 CC Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 CC Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 CC Waterston R., Wilson R.K.;
 CC "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 CC LT2.";
 CC Nature 413:852-856(2001).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AB008754; AAL20162.1; ALT_INIT; Genomic_DNA.
 CC SMR; Q82PZ5; 2-213.
 CC DR StyGene; SG7777; ycfC.
 CC DR HAMAP; MF_00695; -; 1.
 CC DR InterPro; IPR007451; DUF489.
 CC DR Pfam; PF04356; DUF489; 1.
 CC KM Complete proteome; Hypothetical protein.
 CC SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;
 SQ
 Query Match 97.4%; Score 111; DB 1; Length 213;
 Best Local Similarity 95.7%; Pred. No. 1.5e-09;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 NYDITLALAGICQSARLVQOLA 23
 |||||
 Db 4 NYDITLALAGICQSARLVQOLA 26
 |||||

RESULT 9

OSPMJ3_SALPA PRELIMINARY; PRT; 213 AA.

AC 05PMJ3; 01-FEB-2005 (TREMBlrel. 29, Created)
 DT 01-FEB-2005 (TREMBlrel. 29, last sequence update)
 DT 01-FEB-2005 (TREMBlrel. 29, last annotation update)
 DE Hypothetical protein ycfC.
 GN Name=ycfC; OrderedLocustNames=SPR1617;
 OS Salmonella paratyphi-a.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 NCBI_TaxID=54388;
 RX NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 9150;
 RA Mclelland M., Sanderson K.E., Clifton S.W., Latreille P.,
 RA Porrolik S., Sabo A., Meyer R., Bieri T., Ozeraky P., Mclelland M.,
 RA Hartline C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
 RA Kohlberg S., Strong C., Du P., Carter J., Kremliki C., Layman D.,
 RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Mink P., Florea L.,
 RA Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
 RA Spiegh J., Wilson R.K.;
 RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
 RT restricted serovars of Salmonella enterica that cause typhoid.";
 RL Nat. Genet. 36:1268-1274(2004).
 DR EMBL; CP000026; AA07544.1; -; Genomic_DNA.
 DR SMR; Q5PMJ3; 2-213.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 213 AA; 22916 MW; 63192DIC807C5D58 CRC64;

Query Match 97.4%; Score 111; DB 2; Length 213;
 Best local similarity 95.7%; Pred. No. 1.5e-09;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSARLVQQLA 23
 DB 4 NYDITLALSGICQSARLVQQLA 26

RESULT 10
 Q57OC1_SALCH PRELIMINARY; PRT; 215 AA.
 AC 057OC1; 10-MAY-2005 (TREMBlrel. 30, Created)
 DT 10-MAY-2005 (TREMBlrel. 30, last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, last annotation update)
 DE Hypothetical protein yefC.
 GN Name=yefC; OrderedLocustNames=SC1184;
 OS Salmonella cholerae-suis (Salmonella enterica).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 NCBI_TaxID=591;
 RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SC-B67;
 RA PubMed=15781495;
 RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
 RA Wang H.-S., Lee Y.-S.;
 RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
 RT highly invasive and resistant zoonotic pathogen.";
 RL Nucleic Acids Res. 33:1690-1698(2005).
 DR EMBL; AB017220; AA65090.1; -; Genomic_DNA.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 215 AA; 23157 MW; C153820629FA8BD CRC64;

Query Match 93.9%; Score 107; DB 2; Length 215;
 Best local similarity 95.5%; Pred. No. 6.7e-09;

Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSARLVQQL 22
 DB 6 NYDITLALSGICQSARLVQQL 27

RESULT 11
 Q65VY5_MANSW PRELIMINARY; PRT; 204 AA.
 AC 065VY5; 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=MS0298;
 OS Mannheimia succiniciproducens (strain MBEL55E).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Mannheimia.
 NCBI_TaxID=221988;
 RX NUCLEOTIDE SEQUENCE.
 RC PubMed=15378067; DOI=10.1038/nbt1010;
 RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
 RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;
 RT "The genome sequence of the capnophilic rumen bacterium Mannheimia
 RT succiniciproducens.";
 RL Nat. Biotechnol. 22:1275-1281(2004).
 DR EMBL; AB016827; AA036905.1; -; Genomic_DNA.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 204 AA; 23036 MW; 3D64A70A30C33CB CRC64;

Query Match 93.0%; Score 106; DB 2; Length 204;
 Best local similarity 87.0%; Pred. No. 9.3e-09;
 Matches 20; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSARLVQQLA 23
 DB 4 NYDITLALAGVQSARLVQQLA 26

RESULT 12
 Q6D4E8_ERWCT PRELIMINARY; PRT; 212 AA.
 AC 06D4E8; 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=BCA2443;
 OS Brwnia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 NCBI_TaxID=29471;
 RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SCRI 1043 / ATCC BAA-672;
 RA PubMed=15263089; DOI=10.1073/pnas.0402424101;
 RA Bell K.S., Sebatina M., Pritchard L., Holden M.T.G., Hyman L.J.,
 RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
 RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
 RA Fraser A., Hance Z., Haubert H., Jagsels K., Moule S., Norbertczak H.,
 RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
 RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
 RT "Genome sequence of the enterobacterial phytopathogen Brwnia
 RT carotovora subsp. atroseptica and characterization of virulence
 RT factors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
 DR EMBL; BX950851; CAG75345.1; -; Genomic_DNA.
 DR SMR; Q6D4E8; 2-207.
 DR InterPro; IPR007451; DUF489.

```

DR Pfam: PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 212 AA; 22892 MW; A67813005F328150 CRC64;

Query Match
Best Local Similarity 92.1%; Score 105; DB 2; Length 212;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NYDITLALAGICQSARLVQOLA 23
   |||:|||||:|||||:|||||
DB 4 NYEITTLALAGICQSARLVQOLA 26

RESULT 13
Y1850_PASMU STANDARD; PRT; 203 AA.
ID Y1850_PASMU STANDARD; PRT; 203 AA.
AC Q9CXY8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein PM1850.
GN OrderedLocNames=PM1850;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
CX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
RT Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc Natl Acad Sci U S A. 98:3460-3465(2001).

-1- SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AE006223; AK03934.1; -; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 203 AA; 22717 MW; 5B591BFF0E181EDF CRC64;

Query Match
Best Local Similarity 90.4%; Score 103; DB 1; Length 203;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 NYDITLALAGICQSARLVQOLA 23
   |||:|||||:|||||:|||||
DB 3 NYDITLALAGICQSARLVQOLA 25

RESULT 14
Y2805_PHOLL STANDARD; PRT; 208 AA.
ID Y2805_PHOLL STANDARD; PRT; 208 AA.
AC Q7N3B4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein plu2805.
GN OrderedLocNames=plu2805;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
CX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TT01;

```

```

RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchard E., Ruanlok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taouit S., Bocs S., Bouraux-Eude C., Chandelier M., Charles J.-F.,
RA Dassa E., Deroo R., Derzelle S., Freysinet G., Gaudault S.,
RA Medigue C., Lanois A., Powell K., Sigulier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: BX571868; CA815179.1; -; Genomic_DNA.
DR Photobact; plu2805; -.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 208 AA; 22831 MW; 3B484FA1FBFB36AD CRC64;

Query Match
Best Local Similarity 85.1%; Score 97; DB 1; Length 208;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 NYDITLALAGICQSARLVQOLA 23
   |||:|||||:|||||:|||||
DB 4 NYEITTLALAGICQSARLVQOLA 26

RESULT 15
Y638_HABIN STANDARD; PRT; 205 AA.
ID Y638_HABIN STANDARD; PRT; 205 AA.
AC P44796;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein HI0638.
GN OrderedLocNames=HI0638;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
CX NCBI_TaxID=727;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Rd / KM20 / ATCC 51907;
RC MEDLINR=95350630; PubMed=7542800;
RX Fletschmann R.D., Adams M.D., White O., Clayton R.A., Kirsch E.F.,
RX Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RX McKenney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RX Scott J.D., Shirley R., Liu L.-I., Glodde A., Kelley J.M.,
RX Weidman J.F., Phillips C.A., Spriggs T., Hedblom B., Cotton M.D.,
RX Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RX Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RX Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RX Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: U32747; AAC22298.1; -; Genomic_DNA.
DR PIR; I64155; I64155.

```


DR TIGR; H10638; -;
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 23235 MW; 81F31B0186BF82DA CRC64;

Query Match 83.3%; Score 95; DB 1; Length 205;
Best Local Similarity 78.3%; Pred. No. 5.8e-07;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NTYDITLALAGICQSAKLVQOLA 23
||:|||||:|||||:
Db 3 NYHDIYTLALAGVCSAKLVHOLA 25

Search completed: January 23, 2006, 09:31:07
Job time : 61.7 secs

This Page Blank (usps)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838a-3
Perfect score: 114
Sequence: 1 NYDITLALAGICQSARLVQQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA_Main:*
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	23	4	US-10-665-184-3
2	114	100.0	23	5	US-10-942-300-3
3	114	100.0	23	5	US-10-501-838a-3
4	114	100.0	24	4	US-10-665-184-25
5	114	100.0	24	5	US-10-942-300-25
6	114	100.0	25	4	US-10-501-838a-25
7	114	100.0	25	5	US-10-501-838a-19
8	114	100.0	25	5	US-10-501-838a-23
9	114	100.0	26	5	US-10-501-838a-21
10	114	100.0	29	5	US-10-501-838a-18
11	114	100.0	30	4	US-10-665-184-22
12	114	100.0	30	5	US-10-942-300-22
13	114	100.0	30	5	US-10-501-838a-20
14	114	100.0	30	5	US-10-501-838a-22
15	114	100.0	31	4	US-10-665-184-35
16	114	100.0	31	5	US-10-942-300-35
17	114	100.0	31	5	US-10-501-838a-35
18	114	100.0	213	4	US-10-665-184-61
19	114	100.0	213	5	US-10-942-300-61
20	114	100.0	213	5	US-10-501-838a-55
21	114	94.7	23	4	US-10-665-184-26
22	108	94.7	23	5	US-10-942-300-26
23	108	94.7	30	4	US-10-501-838a-26
24	108	94.7	30	5	US-10-665-184-36
25	108	94.7	30	4	US-10-942-300-36
26	108	94.7	30	5	US-10-501-838a-36
27	103	90.4	23	4	US-10-665-184-2

SUMMARIES

ALIGNMENTS

RESULT 1
US-10-665-184-3
; Sequence 3, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-665-184-3

Query Match 100.0%; Score 114; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYDITLALAGICQSARLVQQLA 23
DB 1 NYDITLALAGICQSARLVQQLA 23

RESULT 2
US-10-942-300-3
; Sequence 3, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: Patentin version 3.2
;; SEQ ID NO 3
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-10-942-300-3

Query Match 100.0%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITTLAAGICQSARLVQOLA 23
Db 1 NYDITTLAAGICQSARLVQOLA 23

RESULT 3

US-10-501-838a-3
; Sequence 3, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT FILING DATE: 2004-07-19
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 3
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-501-838a-3

Query Match 100.0%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITTLAAGICQSARLVQOLA 23
Db 1 NYDITTLAAGICQSARLVQOLA 23

RESULT 4

US-10-665-184-25
; Sequence 25, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT FILING DATE: 2003-09-17
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli

US-10-665-184-25

Query Match 100.0%; Score 114; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITTLAAGICQSARLVQOLA 23
Db 2 NYDITTLAAGICQSARLVQOLA 24

RESULT 5

US-10-942-300-25
; Sequence 25, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT FILING DATE: 2004-09-16
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-942-300-25

Query Match 100.0%; Score 114; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITTLAAGICQSARLVQOLA 23
Db 2 NYDITTLAAGICQSARLVQOLA 24

RESULT 6

US-10-501-838a-25
; Sequence 25, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT FILING DATE: 2004-07-19
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-501-838a-25

Query Match 100.0%; Score 114; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSARLVQQLA 23
 |||||
 Db 2 NYDITLALAGICQSARLVQQLA 24

RESULT 7
 US-10-501-838A-19
 ; Sequence 19, Application US/10501838A
 ; Publication No. US20050215478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ben-Sasson, Shmuel A.
 ; APPLICANT: Cohen, Elnat
 ; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
 ; FILE REFERENCE: 24348-501 NATL
 ; CURRENT APPLICATION NUMBER: US/10/501,838A
 ; PRIOR FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: PCT/IB03/00968
 ; PRIOR FILING DATE: 2002-02-07
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 19
 ; LENGTH: 25
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic: penetrating peptide
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (25)..(25)
 ; OTHER INFORMATION: wherein recombinant human insulin is coupled to the penetrating
 ; OTHER INFORMATION: peptide via the glycine residue
 US-10-501-838A-19

Query Match 100.0%; Score 114; DB 5; Length 25;
 Best Local Similarity 100.0%; Pred. No. 4e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSARLVQQLA 23
 |||||
 Db 1 NYDITLALAGICQSARLVQQLA 23

RESULT 8
 US-10-501-838A-23
 ; Sequence 23, Application US/10501838A
 ; Publication No. US20050215478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ben-Sasson, Shmuel A.
 ; APPLICANT: Cohen, Elnat
 ; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
 ; FILE REFERENCE: 24348-501 NATL
 ; CURRENT APPLICATION NUMBER: US/10/501,838A
 ; PRIOR FILING DATE: 2004-07-19
 ; PRIOR APPLICATION NUMBER: PCT/IB03/00968
 ; PRIOR FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: US 60/355,396
 ; PRIOR FILING DATE: 2002-02-07
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 23
 ; LENGTH: 25
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic: penetrating peptide
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (25)..(25)

OTHER INFORMATION: wherein the penetrating peptide is coupled to a linearized
 ; OTHER INFORMATION: insulin receptor, which is in turn coupled to recombinant human
 ; OTHER INFORMATION: insulin
 US-10-501-838A-23

Query Match 100.0%; Score 114; DB 5; Length 25;
 Best Local Similarity 100.0%; Pred. No. 4e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSARLVQQLA 23
 |||||
 Db 1 NYDITLALAGICQSARLVQQLA 23

RESULT 9
 US-10-501-838A-21
 ; Sequence 21, Application US/10501838A
 ; Publication No. US20050215478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ben-Sasson, Shmuel A.
 ; APPLICANT: Cohen, Elnat
 ; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
 ; FILE REFERENCE: 24348-501 NATL
 ; CURRENT APPLICATION NUMBER: US/10/501,838A
 ; PRIOR FILING DATE: 2004-07-19
 ; PRIOR APPLICATION NUMBER: PCT/IB03/00968
 ; PRIOR FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: US 60/355,396
 ; PRIOR FILING DATE: 2002-02-07
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 21
 ; LENGTH: 26
 ; TYPE: PRT
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic: penetrating peptide
 ; NAME/KEY: MISC FEATURE
 ; LOCATION: (26)..(26)
 ; OTHER INFORMATION: wherein heparin is coupled to the penetrating peptide via the
 ; OTHER INFORMATION: free amino group of the lysine residue
 US-10-501-838A-21

Query Match 100.0%; Score 114; DB 5; Length 26;
 Best Local Similarity 100.0%; Pred. No. 4.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYDITLALAGICQSARLVQQLA 23
 |||||
 Db 1 NYDITLALAGICQSARLVQQLA 23

RESULT 10
 US-10-501-838A-18
 ; Sequence 18, Application US/10501838A
 ; Publication No. US20050215478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ben-Sasson, Shmuel A.
 ; APPLICANT: Cohen, Elnat
 ; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
 ; FILE REFERENCE: 24348-501 NATL
 ; CURRENT APPLICATION NUMBER: US/10/501,838A
 ; PRIOR FILING DATE: 2004-07-19
 ; PRIOR APPLICATION NUMBER: PCT/IB03/00968
 ; PRIOR FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: US 60/355,396
 ; PRIOR FILING DATE: 2002-02-07
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 18

LENGTH: 29
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic: penetrating peptide
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (26)-(29)
OTHER INFORMATION: cleavable linker peptide
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (26)-(29)
OTHER INFORMATION: wherein recombinant human insulin is coupled to the penetrating
US-10-501-838A-18

Query Match 100.0%; Score 114; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 4, 6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYDITLALGICOSARLVQOLA 23
Db 1 NYDITLALGICOSARLVQOLA 23

RESULT 11
US-10-665-184-22

Sequence 22, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501CIP
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 22
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Penetrating peptide
FEATURE:
NAME/KEY: MOD RES
LOCATION: (1)-(1)
OTHER INFORMATION: ACETYLATION
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (30)-(30)
OTHER INFORMATION: wherein Xaa is lysine-NH2
US-10-665-184-22

Query Match 100.0%; Score 114; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 4, 8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYDITLALGICOSARLVQOLA 23
Db 1 NYDITLALGICOSARLVQOLA 23

RESULT 12
US-10-942-300-22

Sequence 22, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-503
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 22
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Acylated Penetrating Peptide
NAME/KEY: MISC FEATURE
LOCATION: (30)-(30)
OTHER INFORMATION: wherein Xaa is lysine having a free amino group that is acylated
US-10-942-300-22

Query Match 100.0%; Score 114; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 4, 8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYDITLALGICOSARLVQOLA 23
Db 1 NYDITLALGICOSARLVQOLA 23

RESULT 13
US-10-501-838A-20

Sequence 20, Application US/10501838A
Publication No. US20050215478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: US/10/501,838A
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 20
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic: penetrating peptide
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (26)-(29)
OTHER INFORMATION: cleavable linker peptide
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (30)-(30)
OTHER INFORMATION: wherein heparin is coupled to the penetrating peptide via the
US-10-501-838A-20

Query Match 100.0%; Score 114; DB 5; Length 30;

```
Best Local Similarity    100.0%; Pred. No. 4.8e-11;
Matches      23; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

RESULT 14
US-10-501-838A-22

```

1  / Sequence 22 Application US/10501838A
2  / Publication No. US20050215478A1
3  /
4  / GENERAL INFORMATION:
5  / APPLICANT: Ben-Sasson, Shmuel A.
6  /
7  / TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
8  / TITLE OF INVENTION: Biological Barrier
9  / FILE REFERENCE: 24348-501 NATL
10 / CURRENT APPLICATION NUMBER: US/10/501,838A
11 / CURRENT FILING DATE: 2004-07-19
12 / PRIOR APPLICATION NUMBER: PCT/IB03/00968
13 / PRIOR FILING DATE: 2003-02-07
14 / PRIOR APPLICATION NUMBER: US 60/355,396
15 / PRIOR FILING DATE: 2002-02-07
16 / NUMBER OF SEQ ID NOS: 72
17 / SOFTWARE: PatentIn version 3.2
18 /
19 / SEQ ID NO 22
20 /
21 / LENGTH: 30
22 /
23 / TYPE: PRT
24 / ORGANISM: Artificial sequence
25 /
26 / FEATURE:
27 / OTHER INFORMATION: Synthetic; penetrating peptide
28 /
29 / NAME/KEY: MISC_FEATURE
30 / LOCATION: (27)..(27)
31 /
32 / OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
33 / OTHER INFORMATION: groups of the lysine residue
34 /
35 / FEATURE:
36 / NAME/KEY: MISC_FEATURE
37 / LOCATION: (30)..(30)
38 /
39 / OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
40 / OTHER INFORMATION: groups of the lysine residue
41 /
42 / FEATURE:
43 / NAME/KEY: MISC_FEATURE
44 / LOCATION: (30)..(30)
45 /
46 / OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
47 / OTHER INFORMATION: peptide via the free amino groups of the lysine residue
48 /
49 / US-10-501-838A-22

```

Query Match	100.0%	Score 114	DB 5	Length 30
Best Local Similarity	100.0%	Pred. No. 4.8e-11		
Matches	23	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0

Oy	1 NYYDITLALAGICQSARLVQOLA 23
Dd	1 NYYDITLALAGICQSARLVQOLA 23

RESULT 15
US-10-665-184-35

```

: Sequence 35, Application US/10665184
: Publication No. US20040146549A1
:
: GENERAL INFORMATION:
: APPLICANT: Ben-Saason, Shmuel
: APPLICANT: Cohen, Eliaht
: TITLE OR INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
: TITLE OR INVENTION: Biological Barrier
: FILE REFERENCE: 24348-501CIP
: CURRENT APPLICATION NUMBER: US/10/665,184
: CURRENT FILING DATE: 2003-09-17
: PRIOR APPLICATION NUMBER: PCT/IB03/00968
: PRIOR FILING DATE: 2003-02-07
: PRIOR APPLICATION NUMBER: 60/355,396

```

```

; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35

```

```

? ORGANISM: Artificial
? FEATURE:
? OTHER INFORMATION: Penetrating peptide
?
? FEATURE: NAME/KEY: MOD_RES
? LOCATION: (1)..(1)
? OTHER INFORMATION: ACETYLTATION
?
? FEATURE: NAME/KEY: PEPTIDE
? LOCATION: (31)..(31)
? OTHER INFORMATION: wherein Xaa is Lysal
JS-10-665-184-35

```

Query Match	100.0%	Score 114;	DB 4;	Length 31;
Best Local Similarity	100.0%	Pred. No. 5e-11;		
Matches 23;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

[illegible]

Search completed: January 23, 2006, 11:52:32
Job time : 60 secs

This Page Blank (usphy)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-4

Sequence: 1 A1YDRTIAFGICQALVAVQVVA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1908:*
2: geneseqp1908:*
3: geneseqp2008:*
4: geneseqp2008:*
5: geneseqp2008:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	100.0	23	6 ADB16889	Adb16889 Vibrio ch
2	111	100.0	23	9 AEB08221	Aeb08221 Vibrio ch
3	111	100.0	204	9 AEB08264	Aeb08264 Vibrio ch
4	74	66.7	217	7 ABO61272	Ab061272 Klebsiell
5	73	65.8	22	6 ADB16910	Adb16910 Escherich
6	73	65.8	22	9 AEB08244	Aeb08244 Escherich
7	73	65.8	29	6 ADB16920	Adb16920 Escherich
8	73	65.8	29	9 AEB08254	Aeb08254 Escherich
9	72	64.9	23	6 ADB16887	Adb16887 Escherich
10	72	64.9	23	6 ADB16888	Adb16888 Escherich
11	72	64.9	23	6 ADB16909	Adb16909 Escherich
12	72	64.9	23	9 AEB08220	Aeb08220 Escherich
13	72	64.9	23	9 AEB08219	Aeb08219 Pasteurel
14	72	64.9	23	9 AEB08243	Aeb08243 Escherich
15	72	64.9	24	6 ADB16908	Adb16908 Escherich
16	72	64.9	24	9 AEB08242	Aeb08242 Escherich
17	72	64.9	25	6 ADB16922	Adb16922 Escherich
18	72	64.9	25	6 ADB16906	Adb16906 Escherich
19	72	64.9	26	6 ADB16904	Adb16904 E.coli pe
20	72	64.9	29	6 ADB16921	Adb16921 E.coli pe
21	72	64.9	30	6 ADB16903	Adb16903 E.coli pe
22	72	64.9	30	6 ADB16919	Adb16919 Escherich
23	72	64.9	30	6 ADB16905	Adb16905 Escherich
24	72	64.9	30	9 AEB08239	Aeb08239 Penetrati

25	72	64.9	30	9 AEB08253	Aeb08253 Penetrati
26	72	64.9	31	6 ADB16918	Adb16918 Escherich
27	72	64.9	31	9 AEB08252	Aeb08252 Penetrati
28	72	64.9	203	9 AEB08262	Aeb08262 Pasteurel
29	72	64.9	213	9 AEB08263	Aeb08263 Escherich
30	68	61.3	216	7 ADF07708	Adf07708 Bacterial
31	61	55.0	209	6 ABW67069	Abw67069 Photorhab
32	58	52.3	23	6 ADB16890	Adb16890 Buchnera
33	58	52.3	23	9 AEB08222	Aeb08222 Buchnera
34	58	52.3	211	9 AEB08265	Aeb08265 Buchnera
35	56	50.5	23	6 ADB16886	Adb16886 Haemophil
36	56	50.5	23	6 ADB16912	Adb16912 Haemophil
37	56	50.5	23	9 AEB08218	Aeb08218 Haemophil
38	56	50.5	23	9 AEB08246	Aeb08246 Haemophil
39	56	50.5	29	6 ADB16916	Adb16916 Escherich
40	56	50.5	29	9 AEB08250	Aeb08250 Penetrati
41	56	50.5	205	4 AAM50230	Aam50230 Haemophil
42	56	50.5	205	9 AEB08261	Aeb08261 Haemophil
43	49	44.1	23	6 ADB16892	Adb16892 Xylella f
44	49	44.1	23	9 AEB08224	Aeb08224 Xylella f
45	49	44.1	204	9 AEB08267	Aeb08267 Xylella f

ALIGNMENTS

RESULT 1
ADb16889 standard; peptide, 23 AA.
XX
AC ADB16889;
XX
DT 20-NOV-2003 (first entry)
XX
DE Vibrio cholerae VC1127 penetrating peptide 4.
XX
KW penetrating peptide; epithelial; endothelial; tight junction; diabetes;
KW infertility; hormone; vitamin deficiency; neurodegenerative;
KW cardiovascular; haematological; endocrine disorder; obesity;
KW neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
KW osteopathic; cytoskeletal; nootropic.
XX
OS Vibrio cholerae.
XX
PN WO2003066859-A2.
XX
PD 14-AUG-2003.
XX
PF 07-FEB-2003; 2003WO-1B000968.
XX
PR 07-FEB-2002; 2002US-0355396P.
XX
PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Ben-Sasson SA, Cohen E;
XX
DR WPI; 2003-697452/66.
XX
PT New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
PS Claim 2; Page 14; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, haematological and endocrine disorders, as well as obesity and neoplastic disease. CC Accordingly, the peptides of this invention can be used in compositions CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic, CC cyostatic or neurotropic activities. This peptide is from VC1127 of Vibrio CC cholera and is penetrating peptide 4 of the invention.

Sequence 23 AA;
SQ

Query Match 100.0%; Score 111; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AYDRTTAPAGICQAVLVQOVA 23
1 AYDRTTAPAGICQAVLVQOVA 23

RESULT 2
AEB08221
ID AEB08221 standard; peptide; 23 AA.
XX
AC AEB08221;
XX
DT 25-AUG-2005 (first entry)
XX
DE Vibrio cholerae VC1127 penetrating peptide 4, SEQ ID NO: 4.
XX

Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
antidiabetic; endocrine disease; gastrointestinal disease;
metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
degenerative; musculoskeletal disease; ocular disease; ophthalmological;
neurodegenerative disease; neuroprotective; Alzheimer's disease;
neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
anticonvulsant; genetic disorder; cardiovascular disease;
cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
coronary artery disease; cardiac; vasotropic; obesity; anorectic;
nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
genitourinary disease; hematological disease; antianemic; anemia;
autoimmune disease; immunosuppressive; immune deficiency;
immunostimulant; infectious disease; antimicrobial; infection;
erectile dysfunction; andrology; major depressive disorder;
antidepressant; psychiatric disorder; pain; analgesic;
bacterial infection; antibacterial; viral infection; virocidic;
fungal infection; fungicide; parasitic infection; antiparasitic;
renal failure; antifertility; antirheumatic; cyostatic;
antiinflammatory; hepatotropic; hepatitis B virus infection.

Vibrio cholerae.
OS
PN US2005136103-A1.
XX
PD 23-JUN-2005.
XX
PF 16-SEP-2004; 2004US-00942300.
XX
PR 17-SEP-2003; 2003US-00664989.
XX
PR 17-SEP-2003; 2003US-00665184.
XX
PR 17-SEP-2003; 2003US-0503615P.
XX
PA (BENS)/ BEN-SASSON S A.
XX
PA (COHEN) COHEN E.
XX
PI Ben-Sasson SA, Cohen E;
XX
DK WPI; 2005-444089/45.
XX
PT Composition used for translocating effectors across barrier such as
XX epithelial cells during treatment of e.g. endocrine disorders comprises

effector sequentially coupled with counter ion and hydrophobic agent.

Claim 46, SEQ ID NO 4, 5pp; English.

The present invention relates to a pharmaceutical composition of
penetrating peptides for transcellular delivery of effector. The
invention comprises the effector sequentially coupled with a counter ion
and at least one hydrophobic agent, where the effector is selectively
encapsulated into a complex. The invention is useful for translocating
effectors across a biological barrier such as epithelial cells and
endothelial cells during treatment and prevention of disease or
pathological conditions (including endocrine disorders, diabetes,
infertility, hormone deficiencies, osteoporosis, ophthalmological
disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
coagulable states, coronary disease, cerebrovascular events, metabolic
disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
hematological disorders, anemia of different entities, immunologic and
rheumatologic disorders, autoimmune diseases, immune deficiencies,
infectious diseases, viral infections, bacterial infections, fungal
infections, parasitic infections, neoplastic diseases, multi-factorial
disorders, impotence, chronic pain, depression, different fibrosis states
and short stature) and for mucosal vaccination against anthrax and
hepatitis B. The present sequence is the Vibrio cholerae VC1127
penetrating peptide. This sequence is used in the effective translocation
of aminoglycoside antibiotics and antifungal agents across an epithelial
barrier.

Sequence 23 AA;
SQ

Query Match 100.0%; Score 111; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AYDRTTAPAGICQAVLVQOVA 23
1 AYDRTTAPAGICQAVLVQOVA 23

RESULT 3
AEB08264
ID AEB08264 standard; protein; 204 AA.
XX
AC AEB08264;
XX
DT 25-AUG-2005 (first entry)
XX
DE Vibrio cholerae protein, SEQ ID NO: 62.
XX

Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
antidiabetic; endocrine disease; gastrointestinal disease;
metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
degenerative; musculoskeletal disease; ocular disease; ophthalmological;
neurodegenerative disease; neuroprotective; Alzheimer's disease;
neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
anticonvulsant; genetic disorder; cardiovascular disease;
cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
coronary artery disease; cardiac; vasotropic; obesity; anorectic;
nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
genitourinary disease; hematological disease; antianemic; anemia;
autoimmune disease; immunosuppressive; immune deficiency;
immunostimulant; infectious disease; antimicrobial; infection;
erectile dysfunction; andrology; major depressive disorder;
antidepressant; psychiatric disorder; pain; analgesic;
bacterial infection; antibacterial; viral infection; virocidic;
fungal infection; fungicide; parasitic infection; antiparasitic;
renal failure; antifertility; antirheumatic; cyostatic;
antiinflammatory; hepatotropic; hepatitis B virus infection.

Vibrio cholerae.
OS

PN US2005136103-A1.
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 PF
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 XX WPI, 2005-444089/45.
 DR
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Disclosure; SEQ ID NO 62; 59pp; English.
 PS
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Vibrio cholerae protein
 CC containing penetrating peptide at N-terminal end.
 CC
 XX
 SQ Sequence 204 AA;
 Query Match 100.0%; Score 111; DB 9; Length 204;
 Best Local Similarity 100.0%; Pred. No. 5.3e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AITDRTIAFAGICQAVLVQOVA 23
 DB 4 AITDRTIAFAGICQAVLVQOVA 26
 RESULT 4
 ID ABO61272 standard; protein; 217 AA.
 XX
 AC ABO61272;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 XX Klebsiella pneumoniae polypeptide seqid 7789.
 DE
 XX
 XX Recombinant expression vector; transcription regulatory element;
 KM Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 OS Klebsiella pneumoniae.
 XX
 PN US6610836-B1.

XX 26-AUG-2003.
 PD
 XX 27-JAN-2000; 2000US-00489039.
 PF
 XX 29-JAN-1999; 99US-0117747P.
 PR
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA Breton GL, Osborne M;
 PI
 XX WPI: 2003-895346/82.
 DR N-PSDB; ACH94823.
 DR
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 PT
 XX
 PS Disclosure; SEQ ID NO 7789; 932pp; English.
 CC
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 CC
 XX
 SQ Sequence 217 AA;
 Query Match 66.7%; Score 74; DB 7; Length 217;
 Best Local Similarity 66.7%; Pred. No. 0.00076;
 Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 3 YDRTIAFAGICQAVLVQOVA 23
 DB 10 YDRTIAFAGICQAVLVQOVA 30
 RESULT 5
 ID ADB16910 standard; peptide; 22 AA.
 XX
 AC ADB16910;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 XX Baccharichia coli YCFC penetrating peptide 32.
 DB
 XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 KM infertility; hormone; vitamin deficiency; neurodegenerative;
 KM cardiovascular; haematological; endocrine disorder; obesity;
 KM neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KM osteopathic; cytostatic; nootropic.
 XX
 OS Baccharichia coli.
 XX
 PN WO2003066859-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 07-FEB-2003; 2003WO-1B000968.
 XX
 PR 07-FEB-2002; 2002US-0355396P.
 XX
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 PI Ben-Sasson SA, Cohen E;
 XX
 XX WPI; 2003-697452/66.
 DR
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PT
 XX

PS Claim 2; Page 15; 60pp; English.

CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dargirin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or nootropic activities. This peptide is from YCFC of
 CC *Bacterichia coli* and is penetrating peptide 32 of the invention.

XX SQ Sequence 22 AA;

XX Query Match 65.8%; Score 73; DB 6; Length 22;
 XX Best Local Similarity 63.6%; Pred. No. 9.5e-05;
 XX Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 IYDRTIAFAGICQAVLVQOVA 23
 OY :|||:|||||:|||||:
 DB 1 MYDITTLALGICQSRALVQOLA 22

RESULT 6
 AEB08244
 ID AEB08244 standard; peptide; 22 AA.

XX AC AEB08244;
 XX 25-AUG-2005 (first entry)

DE *Bacterichia coli* YCFC penetrating peptide 32, SEQ ID NO: 27.

XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW nootropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephropathic;
 KW genitourinary disease; hematological disease; anemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virostatic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antiinfectivity; antipneumatic; cyostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.

XX *Bacterichia coli*.
 OS US2005136103-A1.
 XX 23-JUN-2005.
 XX 16-SEP-2004; 2004US-00942300.
 XX 17-SEP-2003; 2003US-00664989.

PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.

XX (BENS/) BEN-SASSON S A.
 XX (COHE/) COHEN E.
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.

XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.

PS Claim 46; SEQ ID NO 27; 59pp; English.

XX The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for transendothelial delivery of effector. The
 XX invention comprises the effector sequentially coupled with a counter ion
 XX and at least one hydrophobic agent, where the effector is selectively
 XX encapsulated into a complex. The invention is useful for translocating
 XX effectors across a biological barrier such as epithelial cells and
 XX endothelial cells during treatment and prevention of disease or
 XX pathological conditions (including endocrine disorders, diabetes,
 XX infertility, hormone deficiencies, osteoporosis, ophthalmological
 XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 XX Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 XX coagulable states, coronary disease, cerebrovascular events, metabolic
 XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 XX hematological disorders, anemia of different entities, immunologic and
 XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
 XX infectious diseases, viral infections, bacterial infections, fungal
 XX infections, parasitic infections, neoplastic diseases, multi-factorial
 XX disorders, impotence, chronic pain, depression, different fibrosis states
 XX and short stature) and for mucosal vaccination against anthrax and
 XX hepatitis B. The present sequence is the *Bacterichia coli* YCFC
 XX penetrating peptide. This sequence is used in the effective translocation
 XX of aminoglycoside antibiotics and antifungal agents across an epithelial
 XX barrier.

XX SQ Sequence 22 AA;

XX Query Match 65.8%; Score 73; DB 9; Length 22;
 XX Best Local Similarity 63.6%; Pred. No. 9.5e-05;
 XX Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 IYDRTIAFAGICQAVLVQOVA 23
 OY :|||:|||||:|||||:
 DB 1 MYDITTLALGICQSRALVQOLA 22

RESULT 7
 ADB16920
 ID ADB16920 standard; peptide; 29 AA.

XX ADB16920;
 XX 20-NOV-2003 (first entry)

DE *Bacterichia coli* peptide 3 coupled to imaging compound linker, IBW-002V3.

XX *Bacterichia coli* peptide 3 coupled to imaging compound linker, IBW-002V3.
 KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiac; antiarteriosclerotic; osteopathic; cyostatic; nootropic;
 KW imaging linker; penetrating peptide; IBW-002V3.

XX Synthetic.
 OS *Bacterichia coli*.
 XX 17-SEP-2003; 2003US-00664989.

XX Key Location/Qualifiers
 XX Modified-site 1

FT Modified-site /note= "N-terminal acetyl"
 PT 29
 FT /note= "C-terminal amide"
 XX
 XX WO2003066859-A2.
 XX
 XX 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI, 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX
 XX Example 3; Page 40; 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, haematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX Accordingly, the peptide of this invention can be used in compositions
 XX that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 XX cytostatic or neurotropic activities. This peptide sequence is IBM-002V3,
 XX which consists of the Escherichia coli penetrating peptide 3 coupled to
 XX the imaging linker peptide used in an exemplification of the invention.
 XX
 XX Sequence 29 AA;
 XX
 XX Query Match 65.8%; Score 73; DB 6; Length 29;
 XX Best Local Similarity 63.6%; Pred. No. 0.00013;
 XX Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 XX
 XX QY 2 IYDRTAFAGICQAVLVQOVA 23
 XX :|||:|||||:|||||:
 XX 1 MYDITLALGICQSAKLVQOVA 22
 XX
 XX RESULT 8
 XX AEB08254
 XX ID AEB08254 standard; peptide; 29 AA.
 XX
 XX AC AEB08254;
 XX
 XX DT 25-AUG-2005 (first entry)
 XX
 XX Penetrating peptide SEQ: 37 used in composition for mucosal vaccination.
 XX
 XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathia;
 XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorders; Huntington's chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;

KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephropathic;
 KW gastrointestinal disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antihypertensive; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX Unidentified.
 XX
 XX Key Location/Qualifiers
 XX
 XX FT Misc-difference 1
 XX FT /note= "N-terminal acylated"
 XX FT Misc-difference 29
 XX FT /note= "Optionally C-terminal amide, optionally the free
 XX amino group of lysine is acylated with a fatty acid"
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHE/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI, 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 63; SEQ ID NO 37; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for transmembrane delivery of effector. The
 XX invention comprises the effector sequentially coupled with a counter ion
 XX and at least one hydrophobic agent, where the effector is selectively
 XX encapsulated into a complex. The invention is useful for translocating
 XX effectors across a biological barrier such as epithelial cells and
 XX endothelial cells during treatment and prevention of disease or
 XX pathological conditions (including endocrine disorders, diabetes,
 XX infertility, hormone deficiencies, osteoporosis, ophthalmological
 XX disorders, neurodegenerative disorders, Huntington's disease,
 XX Parkinson's disease, multiple sclerosis, hypercoagulable states, hypo-
 XX cardiovascular disorders, atherosclerosis, hyperlipidemia, metabolic
 XX coagulable states, coronary disease, cerebrovascular events, renal failure,
 XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 XX hematological disorders, anemia of different entities, immunologic and
 XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
 XX infectious diseases, viral infections, bacterial infections, fungal
 XX infections, parasitic infections, neoplastic diseases, multi-factorial
 XX disorders, impotence, chronic pain, depression, different fibrosis states
 XX and short stature) and for mucosal vaccination against anthrax and
 XX hepatitis B. The present sequence is a penetrating peptide (IBM-002V3)
 XX used in the composition for mucosal vaccination using a counter ion and a
 XX penetrating peptide.
 XX
 XX Sequence 29 AA;
 XX
 XX Query Match 65.8%; Score 73; DB 9; Length 29;
 XX Best Local Similarity 63.6%; Pred. No. 0.00013;

Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 YDRTIAPAGICQAVLVQOVA 23
 :||:|||||:|||||:
 Db 1 YDITLALAGICQASRLVQOLA 22

RESULT 9

ADBI6887
 ID ADBI6887 standard; peptide; 23 AA.

XX ADBI6887;

DT 20-NOV-2003 (first entry)

DE Paeteurella multocida PM1850 penetrating peptide 2.

XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;

XX infertility; hormone; vitamin deficiency; neurodegenerative;

KW cardiovascular; haematological; endocrine disorder; obesity;

KW neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;

KW osteopathic; cytosatic; nootropic.

XX Paeteurella multocida.

OS MO2003066859-A2.

PD 14-AUG-2003.

PF 07-FEB-2003; 2003WO-IB000968.

PR 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Ben-Sasson SA, Cohen E;

XX WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating
 or preventing e.g. endocrine disorders.

PS Claim 2; Page 14; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
 translocating across a biological barrier. Furthermore, it refers to
 methods that use these peptides to facilitate penetration of a
 biologically active effector molecule such as a drug or other therapeutic
 agent across biological barriers e.g. epithelial or endothelial cells
 sealed by tight junctions. This peptide is derived from a bacterial
 toxin, an integral membrane or extracellular protein and can comprise an
 anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 or enzyme. The effector molecule, however, can comprise for example
 insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 factors. The penetrating peptide is useful for the treatment of various
 conditions including diabetes, infertility, hormone and vitamin
 deficiencies, neurodegenerative, cardiovascular, haematological and
 endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 cytosatic or nootropic activities. This peptide is from PM1850 of
 CC Paeteurella multocida and is penetrating peptide 2 of the invention.

XX Sequence 23 AA;

Query Match 64.9%; Score 72; DB 6; Length 23;
 Best Local Similarity 66.7%; Pred. No. 0.00015;
 Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YDRTIAPAGICQAVLVQOVA 23
 ||:|||||:|||||:
 Db 3 YDITLALAGICQASRLVQOPA 23

RESULT 10
 ADBI6888
 ID ADBI6888 standard; peptide; 23 AA.

XX ADBI6888;

DT 20-NOV-2003 (first entry)

DE Escherichia coli YCFC penetrating peptide 3.

XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;

XX infertility; hormone; vitamin deficiency; neurodegenerative;

KW cardiovascular; haematological; endocrine disorder; obesity;

KW neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;

KW osteopathic; cytosatic; nootropic.

XX Escherichia coli.

OS MO2003066859-A2.

PD 14-AUG-2003.

PF 07-FEB-2003; 2003WO-IB000968.

PR 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Ben-Sasson SA, Cohen E;

XX WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating
 or preventing e.g. endocrine disorders.

PS Claim 2; Page 14; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
 translocating across a biological barrier. Furthermore, it refers to
 methods that use these peptides to facilitate penetration of a
 biologically active effector molecule such as a drug or other therapeutic
 agent across biological barriers e.g. epithelial or endothelial cells
 sealed by tight junctions. This peptide is derived from a bacterial
 toxin, an integral membrane or extracellular protein and can comprise an
 anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 or enzyme. The effector molecule, however, can comprise for example
 insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 factors. The penetrating peptide is useful for the treatment of various
 conditions including diabetes, infertility, hormone and vitamin
 deficiencies, neurodegenerative, cardiovascular, haematological and
 endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 cytosatic or nootropic activities. This peptide is from YCFC of
 CC Escherichia coli and is penetrating peptide 3 of the invention.

XX Sequence 23 AA;

Query Match 64.9%; Score 72; DB 6; Length 23;
 Best Local Similarity 66.7%; Pred. No. 0.00015;
 Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDRTIAPAGICQAVLVQOVA 23
 ||:|||||:|||||:
 Db 3 YDITLALAGICQASRLVQOLA 23

RESULT 11
 ADBI6909
 ID ADBI6909 standard; peptide; 23 AA.

XX ADB16909;
 XX 20-NOV-2003 (first entry)
 XX Escherichia coli YCFC penetrating peptide 31.
 DE Escherichia coli YCFC penetrating peptide 31.
 XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 XX infertility; hormone; vitamin deficiency; neurodegenerative;
 XX cardiovascular; haematological; endocrine disorder; obesity;
 XX neoplastic disease; neuroprotective; cardiant; antihypertensive;
 XX osteopathic; cytoskeletal; neurotropic.
 XX Escherichia coli.
 XX WO2003066859-A2.
 XX 14-AUG-2003.
 XX 07-FEB-2003; 2003WO-1B000968.
 XX 07-FEB-2002; 2002US-0355396P.
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2003-697452/66.
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX Claim 2; Page 15; 60pp; English.
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticonagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antihypertensive, osteopathic,
 CC cytoskeletal or neurotropic activities. This peptide is from YCFC of
 CC Escherichia coli and is penetrating peptide 31 of the invention.
 XX Sequence 23 AA;
 SQ
 Query Match 64.9%; Score 72; DB 6; Length 23;
 Best Local Similarity 66.7%; Pred. No. 0.00015; Indels 0; Gaps 0;
 Matches 14; Conservative 3; Mismatches 4;
 QY 3 YDRTIAAGICQAVALVQQA 23
 DB 3 YDRTIAAGICQAVALVQQA 23
 RESULT 12
 AEB08220
 ID AEB08220 standard; peptide; 23 AA.
 AC AEB08220;
 XX 25-AUG-2005 (first entry)
 XX

DE Escherichia coli YCFC penetrating peptide 3, SEQ ID NO: 3.
 XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antihypertensive;
 XX coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephropathic;
 XX genitourinary disease; hematological disease; anemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; viraemic;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antineumatic; cytoskeletal;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX Escherichia coli.
 XX OS
 XX US2005136103-A1.
 XX 23-JUN-2005.
 XX 16-SEP-2004; 2004US-00942300.
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX (BENS/) BEN-SASSON S A.
 XX (COHE/) COHEN E.
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX Claim 53; SEQ ID NO 3; 59pp; English.
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transmembrane delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes;
 CC infertility; hormone deficiencies, osteoporosis, ophthalmological;
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatological disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of anoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX

SQ Sequence 23 AA;
 Query Match 64.9%; Score 72; DB 9; Length 23;
 Best Local Similarity 66.7%; Pred. No. 0.00015;
 Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 3 YDRTAFAGICQAVLVQOVA 23
 3 YDRTALNGVICQAKLVQOFA 23
 Db
 RESULT 13
 AEB08219
 ID AEB08219 standard; peptide; 23 AA.
 AC AEB08219;
 XX
 XX 25-AUG-2005 (first entry)
 DE Pasteurella multocida PM1850 penetrating peptide 2, SEQ ID NO: 2.
 XX
 XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurologic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephropathic;
 KW genitourinary disease; hematological disease; anemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW renal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antineumatic; cycostatic;
 KW antinflammatory; hepatotropic; hepatitis B virus infection.
 KW
 XX Pasteurella multocida.
 OS
 XX US2005136103-A1.
 PV
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 PF
 XX 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS// BEN-SASSON S A.
 PA (COHEN// COHEN E.
 PI Ben-Sasson SA, Cohen E;
 DR WPI; 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 2; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and

CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Pasteurella multocida PM1850
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 SQ Sequence 23 AA;
 Query Match 64.9%; Score 72; DB 9; Length 23;
 Best Local Similarity 66.7%; Pred. No. 0.00015;
 Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 3 YDRTAFAGICQAVLVQOVA 23
 3 YDRTALNGVICQAKLVQOFA 23
 Db
 RESULT 14
 AEB08243
 ID AEB08243 standard; peptide; 23 AA.
 AC AEB08243;
 XX
 XX 25-AUG-2005 (first entry)
 DE Escherichia coli YCPC penetrating peptide 31, SEQ ID NO: 26.
 XX
 XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurologic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephropathic;
 KW genitourinary disease; hematological disease; anemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antineumatic; cycostatic;
 KW antinflammatory; hepatotropic; hepatitis B virus infection.
 KW
 XX Escherichia coli.
 OS
 XX US2005136103-A1.
 PV
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 PF
 XX 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX

This Page Blank (usps)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 Seconds
(without alignments)
114.551 Million cell updates/sec

Title: US-10-501-838A-4
Sequence: 1 AIDRTIAPGICQAVLVQOVA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents, AA:*

- 1: /cgn2_6/ptodata/1/1aa/5-COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/6-COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/H-COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/RTUS-COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/RE-COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	66.7	217	2	US-09-489-039A-7789 Sequence 7789, Ap
2	68	61.3	216	2	US-09-543-681A-7993 Sequence 7993, Ap
3	45.5	41.0	242	2	US-09-248-796A-15477 Sequence 15477, A
4	45	40.5	136	2	US-09-489-039A-9897 Sequence 9897, Ap
5	45	40.5	418	2	US-09-328-352-4525 Sequence 4525, Ap
6	44	39.6	257	2	US-09-830-230A-202 Sequence 202, App
7	44	39.6	310	2	US-09-830-230A-201 Sequence 201, App
8	44	39.6	650	2	US-08-800-291B-4 Sequence 4, Appli
9	43	38.7	97	2	US-09-270-767-54052 Sequence 38835, A
10	43	38.7	97	2	US-08-800-291B-5 Sequence 5, Appli
11	43	38.7	649	2	US-08-800-291B-6 Sequence 5, Appli
12	43	38.7	649	2	US-09-489-039A-9479 Sequence 9479, Ap
13	42	37.8	436	2	US-09-252-991A-18783 Sequence 18783, A
14	42	37.8	436	2	US-09-252-991A-28888 Sequence 28888, A
15	42	37.8	600	2	US-09-302-540-14821 Sequence 14821, A
16	41.5	37.4	341	2	US-09-252-991A-32629 Sequence 32629, A
17	41.5	37.4	688	2	US-09-252-991A-29691 Sequence 29691, A
18	41.5	37.4	49	2	US-09-178-093B-47 Sequence 47, Appli
19	41	36.9	208	2	US-09-270-767-32540 Sequence 32540, A
20	41	36.9	208	2	US-09-270-767-47757 Sequence 47757, A
21	41	36.9	258	2	US-08-303-861-18 Sequence 18, Appli
22	41	36.9	258	2	US-08-303-861-19 Sequence 19, Appli
23	41	36.9	258	2	US-09-213-343-2 Sequence 2, Appli
24	41	36.9	433	2	US-09-252-991A-25146 Sequence 25146, A
25	41	36.9	433	2	US-09-198-452A-479 Sequence 479, App
26	41	36.9	519	2	US-09-438-185A-454 Sequence 454, App
27	41	36.9	519	2	US-09-438-185A-454 Sequence 454, App

28	41	36.9	559	2	US-09-512-260A-3 Sequence 3, Appli
29	40.5	36.5	331	2	US-09-252-991A-21022 Sequence 21022, A
30	40.5	36.5	359	2	US-09-252-991A-28035 Sequence 28035, A
31	40	36.0	1280	2	US-09-252-991A-18001 Sequence 18001, A
32	39.5	35.6	218	2	US-09-489-039A-8572 Sequence 8572, Ap
33	39.5	35.1	35	2	US-09-205-258-756 Sequence 756, App
34	39	35.1	35	2	US-10-004-860-756 Sequence 756, App
35	39	35.1	68	2	US-09-248-796A-26148 Sequence 26148, A
36	39	35.1	78	2	US-09-205-258-761 Sequence 761, App
37	39	35.1	78	2	US-10-004-860-761 Sequence 761, App
38	39	35.1	145	2	US-09-270-767-43660 Sequence 43660, A
39	39	35.1	241	2	US-09-134-000C-3846 Sequence 3846, Ap
40	39	35.1	265	2	US-09-328-352-4347 Sequence 4347, Ap
41	39	35.1	289	2	US-09-328-352-5601 Sequence 5601, Ap
42	39	35.1	312	1	US-09-014-969-177 Sequence 17, Appli
43	39	35.1	429	2	US-09-489-039A-13214 Sequence 13214, A
44	39	35.1	566	2	US-09-252-991A-32371 Sequence 32371, A
45	39	35.1	567	2	US-09-538-092-1365 Sequence 1365, Ap

ALIGNMENTS

RESULT 1
US-09-489-039A-7789
Sequence 7789, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7789
LENGTH: 217
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match Score 74; DB 2; Length 217;
Best Local Similarity 66.7%; Pred. No. 0.00011;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDRTIAPGICQAVLVQOVA 23
DB 10 YDITLAGVCAARLVQOLA 30

RESULT 2
US-09-543-681A-7993
Sequence 7993, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7993
LENGTH: 216
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7993

Query Match Score 68; DB 2; Length 216;

Best Local Similarity 70.0%; Pred. No. 0.0011;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 DRTAFAGICQAVLVQOVA 23
Db 14 DITLALAGICQASRLVQOIA 33

RESULT 3
US-09-248-796A-15477
Sequence 15477, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:

APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
FILE REFERENCE: 107196.132
CURRENT FILING DATE: 1999-02-12
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 15477
LENGTH: 242
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-15477

Query Match
Best Local Similarity 41.0%; Score 45.5; DB 2; Length 242;
Matches 9; Conservative 5; Mismatches 4; Indels 3; Gaps 1;

Qy 2 YDRTAFAGICQAVLVQOV 22
Db 164 YGTIIAFSGVGN---IIQOV 181

RESULT 4
US-09-489-039A-9897
Sequence 9897, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:

APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT FILING DATE: 2000-01-27
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 9897
LENGTH: 136
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9897

Query Match
Best Local Similarity 40.5%; Score 45; DB 2; Length 136;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 2 YDRTAFAGICQAVLVQOV 22
Db 20 IYQTTIHGACATMALEQM 40

RESULT 5
US-09-328-352-4525
Sequence 4525, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4525
LENGTH: 418
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-4525

Query Match
Best Local Similarity 40.5%; Score 45; DB 2; Length 418;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YDRTAFAGICQAVLV 18
Db 234 YDRTIVLVGVSXAYM 250

RESULT 6
US-09-830-230A-202
Sequence 202, Application US/09830230A
Patent No. 6902893
GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Lyme Disease Vaccines
FILE REFERENCE: PB481US
CURRENT FILING DATE: 2001-09-27
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/057,483
PRIOR FILING DATE: 1997-09-03
PRIOR APPLICATION NUMBER: 60/053,344
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/053,377
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/050,359
PRIOR FILING DATE: 1997-06-20
NUMBER OF SEQ ID NOS: 756
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 202
LENGTH: 257
TYPE: PRT
ORGANISM: Homo sapiens
US-09-830-230A-202

Query Match
Best Local Similarity 39.6%; Score 44; DB 2; Length 257;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 YDRTAFAGICQAVLV 18
Db 230 YDFKRFIFIGCQSIAT 245

RESULT 7
US-09-830-230A-201
Sequence 201, Application US/09830230A
Patent No. 6902893
GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Lyme Disease Vaccines
FILE REFERENCE: PB481US
CURRENT FILING DATE: 2001-09-27
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/US98/12718
PRIOR FILING DATE: 1997-09-03

PRIOR APPLICATION NUMBER: 60/053,344
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/053,377
PRIOR FILING DATE: 1997-07-22
PRIOR APPLICATION NUMBER: 60/050,359
PRIOR FILING DATE: 1997-06-20
NUMBER OF SEQ ID NOS: 756
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 201
LENGTH: 310
TYPE: PRT
ORGANISM: Homo sapiens
US-09-830-230A-201

Query Match 39.6%; Score 44; DB 2; Length 310;
Best Local Similarity 43.8%; Pred. No. 19;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 YDRTIAPAGICQAVL 18
DB 283 YDFKTEFICLQSIAT 298

RESULT 8
US-08-800-291B-4
Sequence 4, Application US/08800291B
Patent No. 6153740
GENERAL INFORMATION:
APPLICANT: J.D. Young & C.R. Cass
TITLE OF INVENTION: CDNA ENCODING NUCLEOSIDE TRANSPORTER
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSES: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,291B
FILING DATE: 13-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/499,314
FILING DATE: 7-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Halie, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07254/044MO1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-291B-4

Query Match 39.6%; Score 44; DB 2; Length 650;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 DRTIAPAGICQAVL 19
DB 178 EQVSPAGICVVAL 193

RESULT 9
US-09-270-767-38835
Sequence 38835, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 38835
LENGTH: 97
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-38835

Query Match 38.7%; Score 43; DB 2; Length 97;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 YDRTIAPAGICQAV 17
DB 67 IYKTCSTWGLCMHVA 82

RESULT 10
US-09-270-767-54052
Sequence 54052, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 54052
LENGTH: 97
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-54052

Query Match 38.7%; Score 43; DB 2; Length 97;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 YDRTIAPAGICQAV 17
DB 67 IYKTCSTWGLCMHVA 82

RESULT 11
US-08-800-291B-5
Sequence 5, Application US/08800291B
Patent No. 6153740
GENERAL INFORMATION:
APPLICANT: J.D. Young & C.R. Cass
TITLE OF INVENTION: CDNA ENCODING NUCLEOSIDE TRANSPORTER
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSES: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,291B
FILING DATE: 13-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/499,314
FILING DATE: 7-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07254/044W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 649 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-291B-5

Query Match 38.7%; Score 43; DB 2; Length 649;
Best Local Similarity 43.8%; Pred. No. 65;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 4 DRTAFAGICQAVLV 19
Db 177 EQLVSPAGICVFIAL 192

RESULT 12
US-08-800-291B-6
Sequence 6, Application US/08800291B
Patent No. 6153740
GENERAL INFORMATION:
APPLICANT: J.D. Young & C.E. Case
TITLE OF INVENTION: CDNA ENCODING NUCLEOSIDE TRANSPORTER
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,291B
FILING DATE: 13-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/499,314
FILING DATE: 7-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07254/044W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 649 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-800-291B-6

Query Match 38.7%; Score 43; DB 2; Length 649;
Best Local Similarity 43.8%; Pred. No. 65;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 4 DRTAFAGICQAVLV 19
Db 177 EQLVSPAGICVFIAL 192

RESULT 13
US-09-489-039A-9479
Sequence 9479, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 9479
LENGTH: 338
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9479

Query Match 37.8%; Score 42; DB 2; Length 338;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 6 TIAFAGICQAVLVQVYA 23
Db 319 TIAFAGICQAVLVQVYA 336

RESULT 14
US-09-252-991A-18783
Sequence 18783, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc U. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18783
LENGTH: 436
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18783

Query Match 37.8%; Score 42; DB 2; Length 436;
Best Local Similarity 56.2%; Pred. No. 61;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AIYDRTAFAGICQAV 16
Db 115 AIYDRTAFAGICQAV 130

RESULT 15
US-09-252-991A-28888

This Page Blank (usps)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:56 (Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec)

Title: US-10-501-838a-4

Perfect score: 111

Sequence: 1 AYDRTIAPAGICQAVLVQVVA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications, AA, New.*
1: /cgn2_6/pcdata/2/pubppa/US08_NEW_PUB.pep.*
2: /cgn2_6/pcdata/2/pubppa/US06_NEW_PUB.pep.*
3: /cgn2_6/pcdata/2/pubppa/US07_NEW_PUB.pep.*
4: /cgn2_6/pcdata/2/pubppa/PCT_NEW_PUB.pep.*
5: /cgn2_6/pcdata/2/pubppa/US09_NEW_PUB.pep.*
6: /cgn2_6/pcdata/2/pubppa/US10_NEW_PUB.pep.*
7: /cgn2_6/pcdata/2/pubppa/US11_NEW_PUB.pep.*
8: /cgn2_6/pcdata/2/pubppa/US60_NEW_PUB.pep.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	40.5	1015	6 US-10-957-569-51	Sequence 51, Appl
2	39	35.1	169	6 US-10-467-657-4958	Sequence 4958, Ap
3	39	35.1	409	6 US-10-821-234-1425	Sequence 1425, Ap
4	39	35.1	545	6 US-10-858-730-90	Sequence 90, Appl
5	38	34.2	122	7 US-11-195-459-12	Sequence 12, Appl
6	38	34.2	122	7 US-11-195-459-12	Sequence 19, Appl
7	38	34.2	201	6 US-10-467-657-2982	Sequence 2982, Ap
8	38	34.2	254	7 US-11-103-957-27	Sequence 27, Appl
9	38	34.2	270	6 US-10-467-657-5806	Sequence 5806, Ap
10	37	33.3	20	6 US-10-623-155-503	Sequence 503, App
11	37	33.3	54	6 US-10-467-657-1542	Sequence 1542, Ap
12	37	33.3	111	7 US-11-103-957-51	Sequence 51, Appl
13	37	33.3	288	7 US-11-052-554A-353	Sequence 353, Appl
14	37	33.3	305	6 US-10-520-820-1	Sequence 1, Appl
15	37	33.3	349	7 US-11-143-980-59	Sequence 59, Appl
16	37	33.3	467	6 US-10-467-657-2612	Sequence 2612, Ap
17	37	33.3	630	6 US-10-517-939-42	Sequence 42, Appl
18	37	33.3	1075	7 US-11-174-150-34	Sequence 34, Appl
19	37	33.3	1114	7 US-11-174-150-35	Sequence 35, Appl
20	36	32.4	178	6 US-10-632-150-34	Sequence 34, Appl
21	36	32.4	178	7 US-11-073-457-34	Sequence 34, Appl
22	36	32.4	178	7 US-11-073-457-34	Sequence 34, Appl
23	36	32.4	212	6 US-10-793-626-1628	Sequence 1628, Ap
24	36	32.4	212	6 US-10-793-626-1628	Sequence 1966, Ap
25	36	32.4	842	6 US-10-645-441-2	Sequence 2, Appl

26	36	32.4	919	6 US-10-821-234-1144	Sequence 1144, Ap
27	35	31.5	166	6 US-10-131-826A-8	Sequence 8, Appl
28	35	31.5	183	6 US-10-821-234-1485	Sequence 1485, Ap
29	35	31.5	183	7 US-11-186-284-111	Sequence 111, App
30	35	31.5	183	7 US-11-186-284-113	Sequence 113, App
31	35	31.5	255	6 US-10-467-657-1846	Sequence 1846, App
32	35	31.5	331	7 US-11-016-564-3	Sequence 3, Appl
33	35	31.5	483	7 US-11-113-424-24	Sequence 24, Appl
34	35	31.5	506	6 US-10-641-678-69	Sequence 69, Appl
35	35	31.5	524	7 US-11-113-424-64	Sequence 64, Appl
36	35	31.5	2764	6 US-10-995-561-691	Sequence 691, App
37	35	31.5	2813	6 US-10-995-561-688	Sequence 688, App
38	35	31.5	2919	6 US-10-821-234-1133	Sequence 1133, Ap
39	34.5	31.1	162	6 US-10-467-657-2510	Sequence 2510, Ap
40	34.5	31.1	943	7 US-11-024-959-487	Sequence 487, App
41	34	30.6	42	6 US-10-512-292A-3	Sequence 3, Appl
42	34	30.6	49	6 US-10-467-657-4058	Sequence 4058, App
43	34	30.6	279	6 US-10-873-528-122	Sequence 122, App
44	34	30.6	319	7 US-11-184-005-7	Sequence 7, Appl
45	34	30.6	391	6 US-10-793-626-1236	Sequence 1236, Ap

ALIGNMENTS

RESULT 1
US-10-957-569-51
; Sequence 51, Application US/10957569
; Publication No. US20050246785A1
; GENERAL INFORMATION:
; APPLICANT: COOK, Zhihong et al.
; TITLE OF INVENTION: PROMOTER, PROMOTER CONTROL ELEMENTS, AND USES
; FILE REFERENCE: 2750-1577PUS3
; CURRENT APPLICATION NUMBER: US/10/957,569
; PRIOR FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 10/950,321
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-957-569-51

Query Match 40.5%; Score 45; DB 6; Length 1015;
Best Local Similarity 34.8%; Pred. No. 6.3;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 AYDRTIAPAGICQAVLVQVVA 23
DB 169 ALQDMTILIGVCAFSVLIVGIA 191

RESULT 2
US-10-467-657-4958
; Sequence 4958, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04

SEQ ID NO 4958
LENGTH: 169
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4958

Query Match 35.1%; Score 39; DB 6; Length 169;
Best Local Similarity 39.1%; Pred. No. 9.1;
Matches 9; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 AYDRTIAFAGICQAVLVQQA 23
Db 22 ALFCQTRFVGKCMTFMFAVA 44

RESULT 3
US-10-821-234-1425
Sequence 1425, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmant, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PE SEQ_genes Version 1.0
SEQ ID NO 1425
LENGTH: 409
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1425

Query Match 35.1%; Score 39; DB 6; Length 409;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 RTIAFAGICQAVLVQV 22
Db 177 RLIFLSGLCIAVAVMAV 194

RESULT 4
US-10-858-730-90
Sequence 90, Application US/10858730
Publication No. US20050255568A1
GENERAL INFORMATION:
APPLICANT: Bailey, Richard B.
APPLICANT: Blomquist, Paul
APPLICANT: Doten, Reed
APPLICANT: Driggers, Edward M.
APPLICANT: Madden, Kevin T.
APPLICANT: O'Leary, Jessica
APPLICANT: O'Toole, George
APPLICANT: Trueheart, Joshua
APPLICANT: Walbridge, Michael J.
APPLICANT: Yorgey, Peter S.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
TITLE OF INVENTION: PRODUCTION
FILE REFERENCE: 14184-030001
CURRENT APPLICATION NUMBER: US/10/858,730
CURRENT FILING DATE: 2004-06-01
PRIOR APPLICATION NUMBER: US 60/475,000
PRIOR FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: US 60/551,860
PRIOR FILING DATE: 2004-03-10
NUMBER OF SEQ ID NOS: 364
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 90
LENGTH: 545
TYPE: PRT
ORGANISM: Thermobifida fusca
US-10-858-730-90

Query Match 35.1%; Score 39; DB 6; Length 545;
Best Local Similarity 35.0%; Pred. No. 33;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 IYDRTIAFAGICQAVLVQ 21
Db 156 LYDRTVGIVGIGRGVVAQ 175

RESULT 5
US-11-195-459-12
Sequence 12, Application US/11195459
Publication No. US20050278803A1
GENERAL INFORMATION:
APPLICANT: Sewalt, Vincent
APPLICANT: Hastings, Craig
APPLICANT: Mealey, Robert
APPLICANT: Hantke, Sabine
APPLICANT: Jung, Rudolf
APPLICANT: Eversard, John
APPLICANT: Allen, Stephen
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ALTERING THE DISULFIDE STATUS OF PRO
FILE REFERENCE: 5718-119 (035718/241421)
CURRENT APPLICATION NUMBER: US/11/195,459
CURRENT FILING DATE: 2005-08-02
PRIOR APPLICATION NUMBER: US/10/005,429
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/250,703
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 122
TYPE: PRT
ORGANISM: Zea mays
US-11-195-459-12

Query Match 34.2%; Score 38; DB 7; Length 122;
Best Local Similarity 43.8%; Pred. No. 9.4;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 IYDRTIAFAGICQAVA 17
Db 32 VIDFTAMWCGPCRAIA 47

RESULT 6
US-11-195-459-19
Sequence 19, Application US/11195459
Publication No. US20050278803A1
GENERAL INFORMATION:
APPLICANT: Sewalt, Vincent
APPLICANT: Hastings, Craig
APPLICANT: Mealey, Robert
APPLICANT: Hantke, Sabine
APPLICANT: Jung, Rudolf
APPLICANT: Eversard, John
APPLICANT: Allen, Stephen
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ALTERING THE DISULFIDE STATUS OF PRO
FILE REFERENCE: 5718-119 (035718/241421)
CURRENT APPLICATION NUMBER: US/11/195,459
CURRENT FILING DATE: 2005-08-02
PRIOR APPLICATION NUMBER: US/10/005,429
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/250,703
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 25

SOFTWARE: Patentin version 3.0
SEQ ID NO 19
LENGTH: 122
TYPE: PRT
ORGANISM: Zoa may
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "n" at position 9, 493, 537, 548, 581, and 584 can be an a, c, g,
US-11-195-459-19

Query Match 34.2% Score 38; DB 7; Length 122;
Best Local Similarity 43.8%; Pred. No. 9.4;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 IYDRTIAPAGICQAVA 17
DB 32 VIDPTAAMCGPCRAIA 47

RESULT 7
US-10-467-657-2982
Sequence 2982, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqMin99, version 1.04
SEQ ID NO 2982
LENGTH: 201
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2982

Query Match 34.2% Score 38; DB 6; Length 201;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 TIAFAGICQAV 16
DB 109 TIAFVIGICALV 119

RESULT 8
US-11-103-957-27
Sequence 27, Application US/11103957
Publication No. US20050281847A1
GENERAL INFORMATION:
APPLICANT: Berthet, Francois-Xavier Jacques
APPLICANT: Lobet, Yves
APPLICANT: Poolman, Jan
APPLICANT: Verlaet, Vincent Georges Christian Louis
TITLE OF INVENTION: Vaccine Composition
FILE REFERENCE: B45261
CURRENT APPLICATION NUMBER: US/11/103,957
CURRENT FILING DATE: 2005-04-12
PRIOR APPLICATION NUMBER: US/10/467,534
PRIOR FILING DATE: 2004-02-03
PRIOR APPLICATION NUMBER: PCT/EP02/01356
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: GB 0103169.9
PRIOR FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 108

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 254
TYPE: PRT
ORGANISM: Chlamydia trachomatis
US-11-103-957-27

Query Match 34.2% Score 38; DB 7; Length 254;
Best Local Similarity 41.7%; Pred. No. 21;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IYDRTIAPAGIC 13
DB 61 IFDHVERGIC 72

RESULT 9
US-10-467-657-5806
Sequence 5806, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqMin99, version 1.04
SEQ ID NO 5806
LENGTH: 270
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5806

Query Match 34.2% Score 38; DB 6; Length 270;
Best Local Similarity 41.9%; Pred. No. 22;
Matches 13; Conservative 3; Mismatches 7; Indels 8; Gaps 2;

QY 1 AIYDRTIAPA----GICQAVLV---QOVA 23
DB 130 ALYKRTDAVCVAQSDGICITLQLVNFWDV 160

RESULT 10
US-10-623-155-503
Sequence 503, Application US/10623155
Publication No. US20050261166A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Peckham, David W.
APPLICANT: Reiter, Marc W.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C20
CURRENT APPLICATION NUMBER: US/10/623,155
CURRENT FILING DATE: 2003-07-17
NUMBER OF SEQ ID NOS: 560
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 503
LENGTH: 20
TYPE: PRT
ORGANISM: Homo sapiens
US-10-623-155-503

Query Match 33.3% Score 37; DB 6; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.9;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 DRTIAFAGICQAVVA 17

Db 5 ERTIVKGTCECA 18

RESULT 11

US-10-467-657-1542
; Sequence 1542, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACT Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMn9, version 1.04
; SEQ ID NO 1542
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1542

Query Match 33.3%; Score 37; DB 6; Length 54;

Best Local Similarity 31.8%; Pred. No. 5.7;
Matches 7; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 2 YDRTIAFAGICQAVVA 23

Db 10 VFOTFVSIGICFAVSGILRFA 31

RESULT 12

US-11-103-957-51
; Sequence 51, Application US/11103957
; Publication No. US20050281847A1
; GENERAL INFORMATION:
; APPLICANT: Berthet, Francois-Xavier Jacques
; APPLICANT: Lobet, Yves
; APPLICANT: Poolman, Jan
; APPLICANT: Verlant, Vincent Georges Christien Louis
; TITLE OF INVENTION: Vaccine Composition
; FILE REFERENCE: B45261
; CURRENT APPLICATION NUMBER: US/11/103,957
; CURRENT FILING DATE: 2005-04-12
; PRIOR APPLICATION NUMBER: US/10/467,534
; PRIOR FILING DATE: 2004-02-03
; PRIOR APPLICATION NUMBER: PCT/EP02/01356
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: GB 0103169.9
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-11-103-957-51

Query Match 33.3%; Score 37; DB 7; Length 141;

Best Local Similarity 44.4%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AYDRTIAFAGICQAVVA 18

|||::|||::|||

Db 114 AYKSFASIGIVSFAL 131

RESULT 13

US-11-052-554A-353
; Sequence 353, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 353
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-353

Query Match 33.3%; Score 37; DB 7; Length 288;

Best Local Similarity 41.2%; Pred. No. 35;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 7 IAFAGICQAVVA 23

Db 16 VSFAGIATFASLILPMA 32

RESULT 14

US-10-520-820-1
; Sequence 1, Application US/10520820
; Publication No. US2006000393A1
; GENERAL INFORMATION:
; APPLICANT: MUTHANILIS S.A.
; TITLE OF INVENTION: Pathogenicity determinants which can be used as targets for devel-
; TITLE OF INVENTION: means for preventing and controlling bacterial infections and/or
; FILE REFERENCE: 1621
; CURRENT APPLICATION NUMBER: US/10/520,820
; CURRENT FILING DATE: 2005-01-07
; PRIOR APPLICATION NUMBER: PCT/EP2003/008209
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: FR 0208636
; PRIOR FILING DATE: 2002-07-09
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-520-820-1

Query Match 33.3%; Score 37; DB 6; Length 305;

Best Local Similarity 38.1%; Pred. No. 38;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 3 YDRTIAFAGICQAVVA 23

Db 69 YDAVIDAQLVKSALVTRLA 89

RESULT 15

US-11-143-980-59
; Sequence 59, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:

; APPLICANT: He, Min
 ; APPLICANT: Hucul, John
 ; APPLICANT: Halcil, Bradley A.
 ; APPLICANT: Wegenaar, Melissa M.
 ; APPLICANT: Graziani, Edmund
 ; APPLICANT: Summers, Mia
 ; APPLICANT: Kulowski, Kerry
 ; APPLICANT: Pong, Kevin
 ; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
 ; TITLE OF INVENTION: Polypeptide
 ; FILE REFERENCE: AM-101426US
 ; CURRENT APPLICATION NUMBER: US/11/143,980
 ; CURRENT FILING DATE: 2005-06-03
 ; PRIOR APPLICATION NUMBER: US 60/664,483
 ; PRIOR FILING DATE: 2005-03-23
 ; PRIOR APPLICATION NUMBER: US 60/576,895
 ; PRIOR FILING DATE: 2004-06-03
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 59
 ; LENGTH: 349
 ; TYPE: PRT
 ; ORGANISM: Streptomyces sp.
 ; US-11-143-980-59

Query Match 33.3%; Score 37; DB 7; Length 349;
 Best Local Similarity 50.0%; Pred. No. 44;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 DRTIAFAGICQAVA 17
 DB 268 DRTVSFLWIAQVLA 281

Search completed: January 23, 2006, 11:53:35
 Job time : 5.1 secs

This Page Blank (Left)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-4
Perfect score: 111
Sequence: 1 A1YDRTIAPAGICQAVLVQOVA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR_80:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	111	100.0	205	2	B82237	conserved hypothe
2	72	64.9	208	2	AH0199	conserved hypothe
3	72	64.9	213	2	S19211	ycfc protein - Bsc
4	72	64.9	213	2	D90829	hypotheical prote
5	72	64.9	213	2	B85687	hypotheical prote
6	69	62.2	215	2	AB0647	conserved hypothe
7	58	52.3	211	2	D84960	hypotheical prote
8	56	50.5	205	2	I64155	hypotheical prote
9	52	46.8	551	2	A12364	hypotheical prote
10	52	46.8	690	2	D98318	ABC protein AGR_L
11	49	44.1	204	2	B82680	conserved hypothe
12	45	40.5	1014	2	T04721	Ca2+-transporting
13	45	40.5	1015	2	H84618	probable Ca2+-ATPa
14	44	39.6	310	2	D70139	conserved hypothe
15	44	39.6	321	2	A10007	lipopolysaccharide
16	44	39.6	336	2	AG0105	sugar transport sy
17	44	39.6	400	2	C97623	aspartate aminotra
18	44	39.6	400	2	AC2846	aspartate transam
19	44	39.6	421	2	AF3116	aspartate transam
20	44	39.6	772	2	T30645	hypotheical prote
21	43	38.7	336	2	B96020	probable sugar upc
22	43	38.7	341	2	A13425	ribose transport s
23	43	38.7	705	2	T48044	hypotheical prote
24	42.5	38.3	278	2	B82388	probable maltose o
25	42	37.8	138	2	S75195	ferric uptake regu
26	42	37.8	275	2	B55224	hypotheical prote
27	42	37.8	300	2	D97257	uncharacterized pr
28	42	37.8	346	2	G82961	hypotheical prote
29	42	37.8	346	2	AB0187	probable iron-side

30	42	37.8	359	2	T36050	probable histidin
31	42	37.8	601	2	H69274	probable long-chai
32	42	37.8	847	2	G93843	conserved hypothe
33	41	36.9	195	2	S24228	BHV-1 protein homo
34	41	36.9	378	2	AF3641	histidine-tRNA lig
35	41	36.9	378	2	AB2083	phosphonate metabo
36	41	36.9	513	2	G83725	GMP synthetase gua
37	41	36.9	514	2	AB6547	hypotheical prote
38	41	36.9	514	2	B72076	polymorphic membra
39	41	36.9	602	2	S55188	hypotheical prote
40	41	36.9	698	2	T13769	NAHD2 dehydrogenas
41	41	36.9	1038	2	T30304	protein RanBP7 - A
42	41	36.5	105	2	S72893	hypotheical prote
43	40.5	36.5	223	2	T26763	hypotheical prote
44	40.5	36.5	647	2	S61973	hypotheical prote
45	40.5	36.5	1672	2	C81675	polymorphic membra

ALIGNMENTS

RESULT 1
B82237
conserved hypotheical protein VC1127 [imported] - Vibrio cholerae (strain N16961 serog.
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence revision 20-Aug-2000 #ext_change 09-Jul-2004
C/Accession: B82237
R.Heldelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.D.;
Chardon, D.; Brumfiel, M.D.; Vamathavan, J.; Bagg, S.; Qin, H.; Dragoi, I.; Sellers, I.
1, R.R.; Mekalanos, J.E.; Fanning, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: AB2035; MUID:20406833; PMID:10952301
A/Accession: B82237
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-205 <HEI>
A/Cross-references: UNIPROT:Q9KX93; UNIPARC:UPI00000C2EB4; GB:AE004193; GB:AE003852; NIT
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A/Gene: VC1127
A/Map position: 1

Query Match 100.0%; Score 111; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 A1YDRTIAPAGICQAVLVQOVA 23
DB 4 A1YDRTIAPAGICQAVLVQOVA 26

RESULT 2
AH0199
conserved hypotheical protein YP01637 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence revision 02-Nov-2001 #ext_change 09-Jul-2004
C/Accession: AH0199
R.Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Terriza, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AH0199
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-208 <KUR>
A/Cross-references: UNIPROT:Q8ZF06; UNIPARC:UPI00000CD828; GB:AL590842; PTDN:CAC90459.1
C/Genetics:
A/Gene: YP01637

Query Match 64.9%; Score 72; DB 2; Length 208;

Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 2 YDRRTAFAGICQAVLVQOVA 23

Db 5 IDNRVLALAGVVALQOVROIA 26

RESULT 12

T04721

Ca2+-transporting ATPase (EC 3.6.3.8) ACA2, calmodulin-regulated [validated] - Arabidops

N/Alternate names: plasma membrane-type calcium ATPase; protein F19F18.130

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 31-Dec-2004

C/Accession: T04721; T52352

R/Beyan, M.; Rose, W.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X

submitted to the Protein Sequence Database, March 1999

A/Reference number: Z15382

A/Accession: T04721

A/Molecule type: DNA

A/Residues: 1-1014 <BRV>

A/Cross-references: UNIPROT:O81108; UNIPARC:UPI000000086C; EMBL:AL035605

A/Experimental source: cultivar Columbia; BAC clone F19F18

R/Harper, J.P.; Hong, B.; Hwang, I.; Guo, H.Q.; Stoddard, R.; Huang, J.F.; Palmgren, M.G

J. Biol. Chem. 273, 1099-1106, 1998

A/Title: A novel calmodulin-regulated Ca-2+-ATPase (ACA2) from Arabidopsis with an N-ter

A/Reference number: Z14518; MUID:98086289; PMID:9422775

A/Accession: T52252

A/Status: preliminary; translated from GB/EMBL/DDBU

A/Molecule type: mRNA

A/Residues: 1-1014 <HAR>

A/Cross-references: UNIPARC:UPI000000086C; EMBL:AF025842; PIDN:AAC2697.1

C/Genetics:

A/Gene: ACA2

A/Map position: 4

A/Intons: 58/3; 77/1; 725/3; 778/3; 836/1; 935/3

A/Note: F19F18.130

C/Function:

A/Description: EC 3.6.3.8 [validated, MUID:98086289]

A/Note: Activity is stimulated by calmodulin; localized to an endomembrane system and no

C/Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding dom

C/Keywords: hydrolase

F/650-821/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 40.5%; Score 45; DB 2; Length 1014;

Best Local Similarity 34.8%; Pred. No. 30;

Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 AYDRRTAFAGICQAVLVQOVA 23

Db 168 ALQDMTLMILGVCAVSLIVGIA 190

RESULT 13

H84618

Probable Ca2+-ATPase [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004

C/Accession: H84618

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.

eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A64420; MUID:20083487; PMID:10617197

A/Accession: H84618

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1015 <STO>

A/Cross-references: UNIPROT:O64806; UNIPARC:UPI0000125174; GB:AE002093; NID:96598414; PI

C/Genetics:

A/Gene: At2g22950

A/Map position: 2

C/Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding dom

Query Match 40.5%; Score 45; DB 2; Length 1015;

Best Local Similarity 34.8%; Pred. No. 30;

Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 AYDRRTAFAGICQAVLVQOVA 23

Db 169 ALQDMTLMILGVCAVSLIVGIA 191

RESULT 14

D70139

conserved hypothetical integral membrane protein BB0317 - Lyme disease spirochete

C/Species: Borrelia burgdorferi (lyme disease spirochete)

C/Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C/Accession: D70139

R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White

son, D.; Peterson, J.; Kexlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,

i Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A/Authors: Smith, H.O.; Venter, J.C.

A/Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.

A/Reference number: A70100; MUID:98065943; PMID:9403685

A/Accession: D70139

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-310 <KLB>

A/Cross-references: UNIPROT:O51297; UNIPARC:UPI0000057426; GB:AE001138; GB:AE000783; NID

A/Experimental source: strain B31

Query Match 39.6%; Score 44; DB 2; Length 310;

Best Local Similarity 43.8%; Pred. No. 14;

Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 YDRRTAFAGICQAVLVQOVA 18

Db 283 YDFKXERFLGQSTIAI 298

RESULT 15

A10007

lipopolyaccharide heptosyltransferase-1 [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C/Accession: A10007

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: A10007

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-321 <KUR>

A/Cross-references: UNIPROT:Q8ZJN6; UNIPARC:UPI00000CD660; GB:AL590842; PIDN:CA080923.1;

C/Genetics:

A/Gene: fliC

C/Superfamily: heptosyltransferase I

Query Match 39.6%; Score 44; DB 2; Length 321;

Best Local Similarity 38.1%; Pred. No. 14;

Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 3 YDRRTAFAGICQAVLVQOVA 23

Db 87 YDRVIDAGQIKSAALITRIA 107

Search completed: January 23, 2006, 09:32:49
Job time : 10.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 Seconds
(without alignments)

267.334 Million cell updates/sec

Title: US-10-501-838a-4

Sequence: 1 AYDRITAFAGICQAVLVQOVA 23

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: 1: uniprot_05_80;*
2: uniprot_crembl;* Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	100.0	205	1 Y1127_VIBCH	Q9KX99 vibrio chol
2	105	94.6	205	1 Y1129_VIBPA	Q87Q00 vibrio para
3	101	91.0	205	1 Y1142_VIBBY	Q7MLT5 vibrio vuln
4	101	91.0	205	1 Y2927_VIBBU	Q8DBP4 vibrio vuln
5	87	78.4	205	2 Q583W6_VIBP1	Q5E366 vibrio fisc
6	87	78.4	205	2 Q6L719_PHOPR	Q6L119 photobacter
7	75	67.6	215	1 Y1650_HAEDU	Q9VJ35 haemophilus
8	72	64.9	203	1 Y1850_PASMU	Q9CJY8 pasteurella
9	72	64.9	208	1 Y1637_YBRPR	Q8ZIF6 yersinia pe
10	72	64.9	208	2 Q689Q3_YBRPS	Q669Q3 yersinia ps
11	72	64.9	213	1 YCFC_ECO57	Q8X736 escherichia
12	72	64.9	213	1 YCFC_ECOL6	Q8E1B7 escherichia
13	72	64.9	213	1 YCFC_ECOL1	P25746 escherichia
14	72	64.9	213	1 YCFC_SHIFL	Q831F8 shigella fl
15	69	62.2	204	2 Q6SVT5_MANSM	Q65VY5 manheimia
16	69	62.2	213	1 YCFC_SALTI	Q8Z7H0 salmonella
17	69	62.2	213	1 YCFC_SALTY	Q8ZPZ5 salmonella
18	69	62.2	213	2 Q5PMJ3_SALPA	Q5PMJ3 salmonella
19	67	60.4	212	2 Q6DAB8_ERWCT	Q6DAB8 erwania car
20	66	59.5	205	1 Y2634_SHRON	Q8ED49 shewanella
21	65	58.6	215	2 Q57QCI_SALCH	Q57QCI salmonella
22	61	55.0	208	1 Y2805_PHOHL	Q7H344 photorhabdu
23	58	52.3	211	1 Y262_BUCAT	P57350 buchnera ap
24	56	50.5	205	1 Y638_HAEIN	P44796 haemophilus
25	56	50.5	205	2 Q4QMS9_HAEI8	Q4QMS9 haemophilus
26	53	47.7	217	1 Y243_BUCBP	Q89AM4 buchnera ap
27	52	46.8	251	2 Q8UQI1_AGR75	Q8UQI1 agrobacteri
28	52	46.8	690	2 Q7CS58_AGR75	Q7CS58 agrobacteri
29	50	45.0	202	2 Q60CA8_METCA	Q60CA8 methylcocci
30	49	44.1	204	1 Y1439_XYLFA	Q9PDE0 xylella fas
31	49	44.1	204	1 Y661_XYLFT	Q87DML xylella fas

32	48	43.2	451	2 Q8R6T3_THETN	Q8R6T3 thermotane
33	47	42.3	450	2 Q8R6T4_LACHR	Q8R6T4 lactobacill
34	46	41.4	165	2 Q9DUQ1_VYIRU	Q9DUQ1 soil-borne
35	46	41.4	165	2 Q9Q9P9_VYIRU	Q9Q9P9 soil-borne
36	46	41.4	165	2 Q9Q9Q1_VYIRU	Q9Q9Q1 soil-borne
37	46	41.4	165	2 Q9Q9Q3_VYIRU	Q9Q9Q3 soil-borne
38	46	41.4	165	2 Q9Q9C5_VYIRU	Q9Q9C5 soil-borne
39	46	41.4	165	2 Q779A2_VYIRU	Q779A2 soil-borne
40	46	41.4	424	2 Q89VAV_BRAJA	Q89VAV bradyrhizob
41	46	41.4	1014	2 Q8L8A0_MEDTR	Q8L8A0 medicago tr
42	45.5	41.0	561	2 Q59M26_CANAL	Q59M26 candida alb
43	45.5	41.0	2253	1 PKDR3_HUMAN	Q59M26 homo sapien
44	45.5	41.0	2253	2 Q5THW3_HUMAN	Q5THW3 homo sapien
45	45	40.5	206	1 Y3359_FESGM	Q87Z15 pseudomonas

ALIGNMENTS

RESULT 1
Y1127_VIBCH STANDARD; PRT; 205 AA.
ID Y1127_VIBCH
AC Q9KX99;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein YC1127;
GN OrderedLocustNames=VC1127;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_Taxid=666;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.P., Haft D.H., Hickey B.K., Peterson J.D., Unyap L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Uettermann T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae."
RL Nature 406:477-483(2000)
CC -SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC EMBL: A5004193; AAF94286.1; -; Genomic_DNA.
CC DR TIGR: B82237; B82237.
CC DR TIGR: VC1127; -;
CC DR HAMAP: MF_00695; -; 1.
CC DR InterPro: IPR007451; DUF489.
CC DR Pfam: PF04356; DUF489; 1.
CC KW Complete proteome; Hypothetical protein.
CC SQ SEQUENCE 205 AA; 22771 MW; CCE29F54740848C9 CRC64;
Query Match 100.0%; Score 111; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 9.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AYDRITAFAGICQAVLVQOVA 23
DB 4 AYDRITAFAGICQAVLVQOVA 26
RESULT 2

```

Y1129 VIBPA          STANDARD;          PRT;          205 AA.
ID Y1129 VIBPA
AC O87OM6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein VP1129.
GN OrderedLocustNames=VP1129;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
NCBI_TaxID=670;
RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STPAIN-RIMD 2210633 / Serotype O3:K6;
RX MBLINB=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BA000031; BAC59392.1; -; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 22661 MW; A420AE1FA4A4BA6D3 CRC64;

Query March          94.6%; Score 105; DB 1; Length 205;
Best Local Similarity 95.5%; Pred. No. 9.5e-09;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LYDRTIAFAGICQAVLVQOVA 23
Db 5 LYDRTIAFAGICQAVLVQOVA 26

RESULT 3
Y1342_VIBVY          STANDARD;          PRT;          205 AA.
ID Y1342_VIBVY
AC Q7MLT5;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein VV1342.
GN OrderedLocustNames=VV1342;
OS Vibrio vulnificus (strain VJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
NCBI_TaxID=196600;
RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP PubMed=14656965; DOI=10.1101/gr.1295503;
RX Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-P.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen.";
RL Genome Res. 13:2577-2587(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BA000037; BAC94106.1; ALT_INT; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 22900 MW; 75A52AC3DA988B2 CRC64;

Query March          91.0%; Score 101; DB 1; Length 205;
Best Local Similarity 86.4%; Pred. No. 4.3e-08;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LYDRTIAFAGICQAVLVQOVA 23
Db 5 LYDRTIAFAGICQAVLVQOVA 26

RESULT 4
Y2927_VIBVU          STANDARD;          PRT;          205 AA.
ID Y2927_VIBVU
AC Q8D8P4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein VV12927.
GN OrderedLocustNames=VV12927;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
NCBI_TaxID=672;
RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STPAIN=CMCP6;
RA Rhee U.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Completed genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB016806; AB011259.1; -; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 75A52AC3DA988B2 CRC64;

Query March          91.0%; Score 101; DB 1; Length 205;
Best Local Similarity 86.4%; Pred. No. 4.3e-08;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LYDRTIAFAGICQAVLVQOVA 23
Db 5 LYDRTIAFAGICQAVLVQOVA 26

RESULT 5
O5E3W6_VIBF1          PRELIMINARY;          PRT;          205 AA.
ID O5E3W6_VIBF1
AC O5E3W6;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical cytosolic protein.
GN OrderedLocustNames=VF1785;
OS Vibrio fischeri (strain ATCC 700601 / ES114).

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=312309;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed:1570394; DOI=10.1073/pnas.0409900102;
RA Ruby E.G., Urbanoweki M., Campbell U., Dunn A., Patni M., Gunsalus R.,
RA Locher P., Lupp C., McCann J., Millikan D., Schaefer A., Stebb E.,
RA Stevens A., Vlasick K., Whistler C., Greenberg B.P.;
RT "Complete genome sequence of Vibrio fischeri: a symbiotic bacterium
RT with pathogenic congeners."
RT Proc. Natl. Acad. Sci. U.S.A. 102:3004-3009(2005).
RL EMBL; CP000020; AAM6280.1; -; Genomic_DNA.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 22828 MW; 33D818DFE1C567FD CRC64;

Query Match 78.4%; Score 87; DB 2; Length 205;
Best Local Similarity 72.7%; Pred. No. 8.4e-06;
Matches 16; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 YDRTIAPAGICQAVLVQOVA 23
DB 5 LFDRTIAPAGICQAVLVQOMA 26
:::|||||:::|||||:::

RESULT 6
O6LT19_PHOPR PRELIMINARY; PRT; 205 AA.
ID O6LT19;
AC O6LT19;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein V1342;
GN Name=V1342; OrderedLocusNames=BPRA1146;
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed:1576425; DOI=10.1126/science.1103341;
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N.,
RA Lauro F.M., Campanaro A., Malacrida G., Simionati B., Cannata N.,
RA Romaldi C., Bartlett D.H., Valle G.;
RT "Life at depth: Photobacterium profundum genome sequence and
RT expression analysis."
RT Science 307:1459-1461(2005).
RL EMBL; CR378666; CAG19557.1; -; Genomic_DNA.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome.
SQ SEQUENCE 205 AA; 22842 MW; DCTB4C940F8317A6 CRC64;

Query Match 78.4%; Score 87; DB 2; Length 205;
Best Local Similarity 81.8%; Pred. No. 8.4e-06;
Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 YDRTIAPAGICQAVLVQOVA 23
DB 5 LYDRTIAPAGICQAVLVQOVA 26
:::|||||:::|||||:::

RESULT 7
ID Y1650_HABDU STANDARD; PRT; 215 AA.
AC 07VL35;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein HD1650.
GN OrderedLocusNames=HD1650;

```

```

OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=3500HP / ATCC 700724;
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE017155; AAP6424.1; -; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 215 AA; 23792 MW; 816D40F97794F8C4 CRC64;

Query Match 67.6%; Score 75; DB 1; Length 215;
Best Local Similarity 80.0%; Pred. No. 0.0081;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 DRTIAPAGICQAVLVQOVA 23
DB 7 DRTIAPAGICQAVLVQOVA 26
:::|||||:::|||||:::

RESULT 8
ID Y1850_PASMU STANDARD; PRT; 203 AA.
AC 09CJY8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein PM1850.
GN OrderedLocusNames=PM1850;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
RT "Complete genome sequence of Pasteurella multocida Pm70."
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
RL -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE006223; AAK03934.1; -; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 203 AA; 22717 MW; 5EB91BBF0E181EDF CRC64;

Query Match 64.9%; Score 72; DB 1; Length 203;
Best Local Similarity 66.7%; Pred. No. 0.0024;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

OY 3 YDRTIAPAGICQAVLVQOVA 23
Db 5 YDITLALAGICQSAKLVQOQA 25

RESULT 9
Y1637 YERPE STANDARD; PRT; 208 AA.
ID Y1637 YERPE STANDARD; PRT; 208 AA.
AC 08ZF06; Q8D001;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein YP01637/Y1798/YP1767.
GN OrderedLocustNames=Y1637, Y1798, YP1767;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,
Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
Feltwell T., Hamlin N., Holtz S., Jorgensen K., Kariyeh A.V.,
Leather S., Mouton S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
Stimmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:553-527(2001).
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
avirulent to humans."
RL DNA Res. 11:179-197(2004).
-1 SIMILARITY: Belongs to the UPF0274 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AJ141419; CAC90459.1; -; Genomic DNA.
DR EMBL; AE013782; AAM85366.1; ALT_INIT; Genomic DNA.
DR EMBL; AE017133; AAS61994.1; -; Genomic DNA.
DR PIR; AH0199; AH0199.
DR SMR; Q8ZF06; 2-207.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
FT CONFLICT 1 M -> MIV (in Ref. 2).

SO SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE499723 CRC64;
Query Match 64.9%; Score 72; DB 1; Length 208;
Best Local Similarity 66.7%; Pred. No. 0.0024;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 YDRTIAPAGICQAVLVQOVA 23
Db 6 YDITLALAGICQSAKLVQOQA 26

RESULT 10
Y06903 YERPS PRELIMINARY; PRT; 208 AA.
ID Y06903 YERPS PRELIMINARY; PRT; 208 AA.
AC Q66903;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=YPTB2431;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=1B32953 / Serotype I;
RX PubMed=1538858; DOI=10.1073/pnas.0404012101;
RA Chai N.S.G., Carniel E., Larimer F.W., Lamerlin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Mottin V.L.,
RA Brubaker R.R., Fowler J., Hinebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX363398; CAH21669.1; -; Genomic DNA.
DR SMR; Q66903; 2-207.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SO SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE499723 CRC64;

Query Match 64.9%; Score 72; DB 2; Length 208;
Best Local Similarity 66.7%; Pred. No. 0.0024;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 YDRTIAPAGICQAVLVQOVA 23
Db 6 YDITLALAGICQSAKLVQOQA 26

RESULT 11
YCFC_ECO57 STANDARD; PRT; 213 AA.
ID YCFC_ECO57 STANDARD; PRT; 213 AA.
AC Q8X736;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein YCFC.
GN Name=YCFC; OrderedLocustNames=1861, EC61604;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=8334;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=0157:H7 / BDL933 / ATCC 700927 / EHCC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,

```

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC.
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Tida T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shimagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
CC EMBL, AB005174; AAG5958.1; -; Genomic_DNA.
CC EMBL, BA000007; BAB35027.1; -; Genomic_DNA.
CC PIR, B85687; B85687.
CC PIR, D90829; D90829.
CC SMR, Q8736; 2-213.
CC HAMAP, MF_00695; -; 1.
CC InterPro, IPR007451; DUF489.
CC Pfam, PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC SEQUENCE 213 AA; 22947 MW; E26F9678C344E2 CRC64;
SQ
Query Match 64.9%; Score 72; DB 1; Length 213;
Best Local Similarity 66.7%; Pred. No. 0.0025;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 3 YDRTTAFAGICQAVLVQOVA 23
DB 6 YDITLALAGICQSARLVQOLA 26
-----
RESULT 12
YCFC_ECOL6 STANDARD; PRT; 213 AA.
AC Q8F1B7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein yfcF.
GN Name=yfcF; OrderedLocustNames=c1511;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC.
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raske D., Buckles E.L., Liu S.-R., Boutin P., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not

```

```

CC removed.
-----
CC EMBL, AB016759; AAN79980.1; ALT_INIT; Genomic_DNA.
CC DR SMR, Q8F1B7; 2-213.
CC DR HAMAP, MF_00695; -; 1.
CC InterPro, IPR007451; DUF489.
CC Pfam, PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC SEQUENCE 213 AA; 22934 MW; 71C8F96F8BA37BB CRC64;
SQ
Query Match 64.9%; Score 72; DB 1; Length 213;
Best Local Similarity 66.7%; Pred. No. 0.0025;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 3 YDRTTAFAGICQAVLVQOVA 23
DB 6 YDITLALAGICQSARLVQOLA 26
-----
RESULT 13
YCFC_ECOL1 STANDARD; PRT; 213 AA.
AC P25746;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein yfcF (ORF-23).
GN Name=yfcF; OrderedLocustNames=b1132;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KL2;
RA Green S.M., Drablie W.T.;
RL Submitted (MAY-1991) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KL2;
RX MEDLINE=92104952; PubMed=1729205;
RA He B., Smith J.M., Zalkin H.;
RT "Escherichia coli purB gene: cloning, nucleotide sequence, and
RT regulation by purK."
RL J. Bacteriol. 174:130-136(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KL2 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1233-1238(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KL2;
RX MEDLINE=97061202; PubMed=8905232;
RA Oehlman T., Alts H., Baba T., Fujita K., Hayaishi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitegawa M., Makino K., Maeda S., Miki T., Mirobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horinouchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to
CC frame shifts.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X59307; CA441995.1; -; Genomic DNA.
DR EMBL; M74924; -; NOT ANNOTATED CDS; Genomic_DNA.
DR EMBL; U00096; AAC74216.1; -; Genomic DNA.
DR EMBL; D90748; BAA35954.1; -; Genomic DNA.
DR EMBL; D90749; BAA35963.1; -; Genomic DNA.
DR PIR; S19211; S19211.
DR PDB; 1QZ4; X-ray; A=2-213.
DR PDB; 1SDI; X-ray; A=2-213.
DR ECHOBASE; EB1321; -.
DR EcoGene; BG11345; ycfC.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM 3D-structure; Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22948 MW; E26EF9698C3CBE42 CRC64;

Query Match 64.9%; Score 72; DB 1; Length 213;
Best Local Similarity 66.7%; Pred. No. 0.0025;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YDRTIAPAGICQAVLVQOVA 23
Db 6 YDITLALAGVQSARLVQOLA 26

RESULT 14
YCF_C SHIFL STANDARD; PRT; 213 AA.
AC Q83LF8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name:ycfC; OrderedLocustNames=SF1151, S1234;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin O., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Ou D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Berna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
```

```
CC removed.
CC -----
DR EMBL; AE005674; AAN42768.1; ALT_INIT; Genomic DNA.
DR EMBL; AE016982; AAP1657.1; -; Genomic DNA.
DR SMR; Q83LF8; 2-213.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22894 MW; 0B0BC698C29FB4D CRC64;

Query Match 64.9%; Score 72; DB 1; Length 213;
Best Local Similarity 66.7%; Pred. No. 0.0025;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YDRTIAPAGICQAVLVQOVA 23
Db 6 YDITLALAGVQSARLVQOLA 26

RESULT 15
Q65VV5 MANSM PRELIMINARY; PRT; 204 AA.
ID Q65VV5;
AC Q65VV5;
DT 25-OCT-2004 (TRENBLREL. 28, Created)
DT 25-OCT-2004 (TRENBLREL. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLREL. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MS0298;
OS Mannheimia succiniciproducens (strain MBR55B).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=221988;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15378067; DOI=10.1038/nbt1010;
RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;
RT "The genome sequence of the capnophilic rumen bacterium Mannheimia
RT succiniciproducens."
RL Nat. Biotechnol. 22:1275-1281(2004).
DR EMBL; AB016827; AAU36905.1; -; Genomic DNA.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 204 AA; 23036 MW; 3D64A70A30C3CCB CRC64;

Query Match 62.2%; Score 69; DB 2; Length 204;
Best Local Similarity 61.9%; Pred. No. 0.0074;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 YDRTIAPAGICQAVLVQOVA 23
Db 6 YDITLALAGVQSARLVQOVA 26
```

Search completed: January 23, 2006, 09:31:08
Job time : 61.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838a-4

Perfect score: 111

Sequence: 1 AAYDRTAFAGICQAVLVQOVA 23

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	100.0	23	4	US-10-665-184-4
2	111	100.0	23	5	US-10-942-300-4
3	111	100.0	23	5	US-10-501-838a-4
4	111	100.0	204	4	US-10-665-184-62
5	111	100.0	204	5	US-10-942-300-62
6	111	100.0	204	5	US-10-501-838a-56
7	73	65.8	22	4	US-10-665-184-27
8	73	65.8	22	5	US-10-942-300-27
9	73	65.8	22	5	US-10-501-838a-27
10	73	65.8	29	4	US-10-665-184-37
11	73	65.8	29	5	US-10-942-300-37
12	73	65.8	29	5	US-10-501-838a-37
13	72	64.9	23	4	US-10-665-184-3
14	72	64.9	23	4	US-10-665-184-3
15	72	64.9	23	4	US-10-665-184-26
16	72	64.9	23	5	US-10-942-300-2
17	72	64.9	23	5	US-10-942-300-3
18	72	64.9	23	5	US-10-942-300-26
19	72	64.9	23	5	US-10-501-838a-2
20	72	64.9	23	5	US-10-501-838a-3
21	72	64.9	23	5	US-10-501-838a-26
22	72	64.9	24	4	US-10-665-184-25
23	72	64.9	24	5	US-10-942-300-25
24	72	64.9	24	5	US-10-501-838a-25
25	72	64.9	25	5	US-10-501-838a-19
26	72	64.9	25	5	US-10-501-838a-23
27	72	64.9	26	5	US-10-501-838a-21

28	72	64.9	29	5	US-10-501-838a-18	Sequence 18, Appl
29	72	64.9	30	4	US-10-665-184-22	Sequence 22, Appl
30	72	64.9	30	4	US-10-665-184-36	Sequence 36, Appl
31	72	64.9	30	5	US-10-942-300-22	Sequence 22, Appl
32	72	64.9	30	5	US-10-942-300-36	Sequence 36, Appl
33	72	64.9	30	5	US-10-501-838a-20	Sequence 20, Appl
34	72	64.9	30	5	US-10-501-838a-22	Sequence 22, Appl
35	72	64.9	30	5	US-10-501-838a-36	Sequence 36, Appl
36	72	64.9	31	4	US-10-665-184-35	Sequence 35, Appl
37	72	64.9	31	5	US-10-942-300-35	Sequence 35, Appl
38	72	64.9	31	5	US-10-501-838a-35	Sequence 35, Appl
39	72	64.9	203	4	US-10-665-184-60	Sequence 60, Appl
40	72	64.9	203	5	US-10-942-300-60	Sequence 60, Appl
41	72	64.9	203	5	US-10-501-838a-54	Sequence 54, Appl
42	72	64.9	213	4	US-10-665-184-61	Sequence 61, Appl
43	72	64.9	213	5	US-10-942-300-61	Sequence 61, Appl
44	72	64.9	213	5	US-10-501-838a-55	Sequence 55, Appl
45	58	52.3	23	4	US-10-665-184-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-10-665-184-4 ; Sequence 4, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across ;
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-665-184-4

Query Match 100.0%; Score 111; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAYDRTAFAGICQAVLVQOVA 23
DB 1 AAYDRTAFAGICQAVLVQOVA 23

RESULT 2
US-10-942-300-4 ; Sequence 4, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a ;
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: Patentin version 3.2
;; SEQ ID NO 4
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Vibrio cholerae
US-10-942-300-4

Query Match 100.0%; Score 111; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 5,6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAYDRTIAFAGICQAVLVQOVA 23
Db 1 AAYDRTIAFAGICQAVLVQOVA 23

RESULT 3
US-10-501-838A-4
; Sequence 4, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elinat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-501-838A-4

Query Match 100.0%; Score 111; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 5,6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAYDRTIAFAGICQAVLVQOVA 23
Db 1 AAYDRTIAFAGICQAVLVQOVA 23

RESULT 4
US-10-665-184-62
; Sequence 62, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elinat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 62
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Vibrio cholerae

US-10-665-184-62

Query Match 100.0%; Score 111; DB 4; Length 204;
Best Local Similarity 100.0%; Pred. No. 5,8e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAYDRTIAFAGICQAVLVQOVA 23
Db 4 AAYDRTIAFAGICQAVLVQOVA 26

RESULT 5
US-10-942-300-62

; Sequence 62, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elinat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 62
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-942-300-62

Query Match 100.0%; Score 111; DB 5; Length 204;
Best Local Similarity 100.0%; Pred. No. 5,8e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAYDRTIAFAGICQAVLVQOVA 23
Db 4 AAYDRTIAFAGICQAVLVQOVA 26

RESULT 6
US-10-501-838A-56

; Sequence 56, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elinat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 56
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-501-838A-56

Query Match 100.0%; Score 111; DB 5; Length 204;
Best Local Similarity 100.0%; Pred. No. 5,8e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY		1 A I Y D R T I A F A G I C Q A V A L V Q Q V A 23
D b		4 A I Y D R T I A F A G I C Q A V A L V Q Q V A 26

```

RESULT 7
US-10-665-184-27
Sequence 27, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
TITLE OF INVENTION: Biological Barrier
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 27
LENGTH: 22
TYPE: PRT
ORGANISM: Escherichia coli
US-10-665-184-27

```

Query Match	65.8%	Score 73	DB 4	Length 22
Best Local Similarity	63.6%	Pred. No. 9.2e-05		
Matches 14; Conservative	4	Mismatches 4	Indels 0	Gaps 0

```

Oy      2  IYDRTIAFAGICQAVALLVQOVA 23
         :|||:|||||:|||||:|
Db      1  MYDITLALAGICQASARLVQOLA 22

```

```

      RESULT 8
US-10-942-300-27
/ Sequence 27, Application US/10942300
/ Publication No. US20050136103A1
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Binat
/ APPLICANT: Ben-Sasson, Shmuel
/ TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
/ TITLE OF INVENTION: Biological Barrier
/ FILE REFERENCE: 24348-503
/ CURRENT APPLICATION NUMBER: US/10/942,300
/ CURRENT FILING DATE: 2004-09-16
/ PRIOR APPLICATION NUMBER: 10/665,164
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 10/664,989
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 60/503,615
/ PRIOR FILING DATE: 2003-09-17
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 27
/ LENGTH: 22
/ TYPE: prt
/ ORGANISM: Escherichia coli
US-10-942-300-27

```

Query Match	65.8%	Score 73	DB 5	Length 22
Best Local Similarity	63.6%	Pred. No. 9.2e-05		
Matches 14, Conservative	4	Mismatches 4	Indels 0	Gaps 0

```

OY      2  IYDRTIAFAGICQAVALLVQQA 23
         :||:|||||:||||:|
Db      1  MYDITLALAGICQSARLVQQLA 22

```

RESULT 9
US-10-501-838A-27
; Sequence 27, Application US/10501838A
; Publication No. US20050215478A1

Query Match	65.8%;	Score 73;	DB 5;	Length 22;
Best Local Similarity	63.6%;	Pred. No. 9.2e-05;		
Matches 14;	Conservative 4;	Mismatches 4;	Indels 0;	Gaps 0;

QY 2 IYDRTIAPAGICQAVLVQQA 23
: || | : | | | | : | | | : |
Db 1 MYDITLALAGICQSARLVQQLA 22

```

10 RESULT 10
11 US-10-665-184-37
12 / Sequence 37, Application US/10665184
13 / Publication No. US20040146549A1
14 /
15 / GENERAL INFORMATION:
16 / APPLICANT: Ben-Sasson, Shmuel
17 / APPLICANT: Cohen, Eilat
18 / TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
19 / TITLE OF INVENTION: Biological Barrier
20 / FILE REFERENCE: 24348-501CIP
21 / CURRENT APPLICATION NUMBER: US/10/665,184
22 / PRIOR FILING DATE: 2003-09-17
23 / PRIOR APPLICATION NUMBER: PCT/IB03/00968
24 / PRIOR FILING DATE: 2003-02-07
25 / PRIOR APPLICATION NUMBER: 60/355,396
26 / PRIOR FILING DATE: 2002-02-07
27 / NUMBER OF SEQ ID NOS: 65
28 / SOFTWARE: PatentIn version 3.2
29 / SEQ ID NO 37
30 / LENGTH: 29
31 / TYPE: PR1
32 / ORGANISM: Artificial
33 / FEATURE:
34 / OTHER INFORMATION: Penetrating peptide
35 /
36 / FEATURE:
37 / NAME/KEY: MOD_RES
38 / LOCATION: (1)..(1)
39 / OTHER INFORMATION: ACETYLATION
40 / FEATURE:
41 / NAME/KEY: PEPTIDE
42 / LOCATION: (29)..(29)
43 / OTHER INFORMATION: wherein Xaa is Lysine-NH2
44 US-10-665-184-37

```

Query Match	65.8%	Score 73	DB 4	Length 29
Best Local Similarity	63.6%	Pred. No. 0.00012		
Matches 14; Conservative	4	Mismatches 4	Indels 0	Gaps 0

QY 2 IYDRTIAFAGICQAVALLVQQA 23

Db 1 MYDITLALAGICQSRALVQQLA 22

RESULT 11

US-10-942-300-37

Sequence 37, Application US/10942300

Publication No. US20050136103A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel

APPLICANT: Cohen, Elina

TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-503

CURRENT APPLICATION NUMBER: US/10/942,300

CURRENT FILING DATE: 2004-09-16

PRIOR APPLICATION NUMBER: 10/665,184

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 10/664,989

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 60/503,615

PRIOR FILING DATE: 2003-09-17

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.2

SEQ ID NO 37

LENGTH: 29

TYPE: PRT

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Acylated Penetrating Peptide

NAME/KEY: MISC FEATURE

LOCATION: (129)-(129)

OTHER INFORMATION: wherein Xaa is lysine having a free amino group that is acylated

US-10-942-300-37

Query Match Best Local Similarity 65.8%; Score 73; DB 5; Length 29; Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 YDRTIAFAGICQAVLVQVQA 23

Db 1 MYDITLALAGICQSRALVQQLA 22

RESULT 12

US-10-501-838A-37

Sequence 37, Application US/10501838A

Publication No. US20050215478A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel A.

APPLICANT: Cohen, Elina

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-501 NATL

CURRENT APPLICATION NUMBER: US/10/501,838A

CURRENT FILING DATE: 2004-07-19

PRIOR APPLICATION NUMBER: PCT/IB03/00968

PRIOR FILING DATE: 2003-02-07

PRIOR APPLICATION NUMBER: US 60/355,396

PRIOR FILING DATE: 2002-02-07

NUMBER OF SEQ ID NOS: 72

SOFTWARE: PatentIn version 3.2

SEQ ID NO 37

LENGTH: 29

TYPE: PRT

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: Synthetic; penetrating peptide

NAME/KEY: MISC FEATURE

LOCATION: (26)-(26)

OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino

OTHER INFORMATION: groups of the lysine residue

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (29)-(29)

OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino

OTHER INFORMATION: groups of the lysine residue

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (29)-(29)

OTHER INFORMATION: wherein another molecule can be coupled to the penetrating

OTHER INFORMATION: peptide via the free amino groups of the lysine residue

US-10-501-838A-37

Query Match Best Local Similarity 65.8%; Score 73; DB 5; Length 29; Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 YDRTIAFAGICQAVLVQVQA 23

Db 1 MYDITLALAGICQSRALVQQLA 22

RESULT 13

US-10-665-184-2

Sequence 2, Application US/10665184

Publication No. US20040146549A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel

APPLICANT: Cohen, Elina

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-501CIP

CURRENT APPLICATION NUMBER: US/10/665,184

CURRENT FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: PCT/IB03/00968

PRIOR FILING DATE: 2003-02-07

PRIOR APPLICATION NUMBER: 60/355,396

PRIOR FILING DATE: 2002-02-07

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.2

SEQ ID NO 2

LENGTH: 23

TYPE: PRT

ORGANISM: Pasteurella multocida

US-10-665-184-2

Query Match Best Local Similarity 64.9%; Score 72; DB 4; Length 23; Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 YDRTIAFAGICQAVLVQVQA 23

Db 3 YDITLALAGVCAAKLVQVQA 23

RESULT 14

US-10-665-184-3

Sequence 3, Application US/10665184

Publication No. US20040146549A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel

APPLICANT: Cohen, Elina

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-501CIP

CURRENT APPLICATION NUMBER: US/10/665,184

CURRENT FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: PCT/IB03/00968

PRIOR FILING DATE: 2003-02-07

PRIOR APPLICATION NUMBER: 60/355,396

PRIOR FILING DATE: 2002-02-07

NUMBER OF SEQ ID NOS: 65

/ SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 3
 / LENGTH: 23
 / TYPE: PRT
 / ORGANISM: Escherichia coli
 US-10-665-184-3

Query Match 64.9%; Score 72; DB 4; Length 23;
 Best Local Similarity 66.7%; Pred. No. 0.00014;
 Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDRTIAPAGICQAVLVQOVA 23
 DB 3 YDITLALAGICQARLVQOLA 23

RESULT 15
 US-10-665-184-26
 / Sequence 26, Application US/10665184
 / Publication No. US20040146549A1
 / GENERAL INFORMATION:
 / APPLICANT: Ben-Sasson, Shmuel
 / TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
 / FILE REFERENCE: 24348-501CIP
 / CURRENT APPLICATION NUMBER: US/10/665,184
 / PRIOR FILING DATE: 2003-09-17
 / PRIOR APPLICATION NUMBER: PCT/IB03/00968
 / PRIOR FILING DATE: 2003-02-07
 / PRIOR APPLICATION NUMBER: 60/355,396
 / PRIOR FILING DATE: 2002-02-07
 / NUMBER OF SEQ ID NOS: 65
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 26
 / LENGTH: 23
 / TYPE: PRT
 / ORGANISM: Escherichia coli
 US-10-665-184-26

Query Match 64.9%; Score 72; DB 4; Length 23;
 Best Local Similarity 66.7%; Pred. No. 0.00014;
 Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDRTIAPAGICQAVLVQOVA 23
 DB 3 YDITLALAGICQARLVQOLA 23

Search completed: January 23, 2006, 11:52:32
 Job time : 60 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-5

Perfect score: 112

Sequence: 1 KIHLLTSLAGICQSAHLVQQLA 23

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq.21.*
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003b:*
8: geneseqp2004a:*
9: geneseqp2005a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	23	6	ADBI6890 Buchnera
2	112	100.0	23	9	AEBO8222 Buchnera
3	112	100.0	211	9	AEBO8265 Buchnera
4	80	71.4	22	6	ADBI6910 Escherich
5	80	71.4	22	6	AEBO8244 Escherich
6	80	71.4	23	6	ADBI6888 Escherich
7	80	71.4	23	6	ADBI6909 Escherich
8	80	71.4	23	6	AEBO8220 Escherich
9	80	71.4	23	9	AEBO8243 Escherich
10	80	71.4	24	6	ADBI6908 Escherich
11	80	71.4	24	6	AEBO8242 Escherich
12	80	71.4	25	6	ADBI6922 Escherich
13	80	71.4	25	6	ADBI6906 Escherich
14	80	71.4	26	6	ADBI6904 Escherich
15	80	71.4	26	6	ADBI6920 Escherich
16	80	71.4	29	6	ADBI6921 Escherich
17	80	71.4	29	6	AEBO8254 Escherich
18	80	71.4	30	6	ADBI6903 Escherich
19	80	71.4	30	6	ADBI6919 Escherich
20	80	71.4	30	6	ADBI6905 Escherich
21	80	71.4	30	9	AEBO8239 Escherich
22	80	71.4	30	9	AEBO8253 Escherich
23	80	71.4	31	6	ADBI6918 Escherich
24	80	71.4	31	9	AEBO8252 Escherich

25	80	71.4	213	9	AEBO8263 Escherich
26	76	67.9	217	7	ABO61272 Klebsiell
27	73	65.2	23	6	ADBI6912 Haemophil
28	73	65.2	23	9	AEBO8246 Haemophil
29	73	65.2	29	6	ADBI6916 Escherich
30	73	65.2	29	9	AEBO8250 Penetrati
31	73	65.2	209	6	ABW67069 Photorhab
32	72	64.3	23	6	ADBI6886 Haemophil
33	72	64.3	23	9	AEBO8218 Haemophil
34	72	64.3	205	4	AAW50230 Haemophil
35	72	64.3	205	9	AEBO8261 Haemophil
36	72	64.3	216	7	ADF07708 Bacteri
37	71	63.4	23	6	ADBI6887 Pasteurel
38	71	63.4	23	9	AEBO8219 Pasteurel
39	71	63.4	203	9	AEBO8262 Pasteurel
40	58	51.8	23	6	ADBI6889 Vibrio ch
41	58	51.8	23	9	AEBO8221 Vibrio ch
42	58	51.8	204	9	AEBO8264 Vibrio ch
43	51	45.5	463	8	ADT55783 Plant pol
44	49	43.8	426	9	ADY34386 Streptomy
45	46	41.1	61	4	ABG29025 Novel hum

ALIGNMENTS

RESULT 1
ADBI6890 standard; peptide; 23 AA.

ADBI6890;

20-NOV-2003 (first entry)

Buchnera aphidicola BU262 penetrating peptide 5.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;

infertility; hormone; vitamin deficiency; neurodegenerative;

cardiovascular; haematological; endocrine disorder; obesity;

neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;

osteopathic; cytoskeletal; nootropic.

Buchnera aphidicola.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-1B000968.

07-FEB-2002; 2002US-035396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating

or preventing e.g. endocrine disorders.

Claim 2; Page 14; 60pp; English.

This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antihypertensive, osteopathic,
 CC cytosolic or neurotropic activities. This peptide is from BU262 of
 CC buchnera aphidicola and is penetrating peptide 5 of the invention.
 CC
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 112; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 4,1e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KIHITITSLAGICQSARLVQOLA 23
 Db 1 KIHITITSLAGICQSARLVQOLA 23
 RESULT 2
 AEB08222 standard; peptide; 23 AA.
 AC AEB08222;
 DT 25-AUG-2005 (first entry)
 DE Buchnera aphidicola BU262 penetrating peptide 5, SEQ ID NO: 5.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurologic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; anitameic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; viruscide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Buchnera aphidicola.
 PN US2005136103-A1.
 PD 23-JUN-2005.
 PF 16-SEP-2004; 2004US-00942300.
 PR 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Saeson SA, Cohen E;
 DR WPI; 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 5, 5pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transseptal delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC parkinson's disease, multiple sclerosis, huntington's disease, hypo-
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Buchnera aphidicola BU262
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 CC
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 112; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 4,1e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KIHITITSLAGICQSARLVQOLA 23
 Db 1 KIHITITSLAGICQSARLVQOLA 23
 RESULT 3
 AEB08265 standard; protein; 211 AA.
 AC AEB08265;
 DT 25-AUG-2005 (first entry)
 DE Buchnera aphidicola protein, SEQ ID NO: 63.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurologic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; anitameic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; viruscide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Buchnera aphidicola.

PN US2005136103-A1.
 XX
 XX 23-JUN-2005.
 PD
 XX
 XX 16-SEP-2004; 2004US-00942300.
 PP
 XX 17-SEP-2003; 2003US-00664989.
 XX
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 P1 WPI, 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Disclousure; SEQ ID NO 63; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Buchnera aphidicola protein
 CC containing penetrating peptide at N-terminal end.
 XX
 XX Sequence 211 AA;
 SQ
 Query Match 100.0%; Score 112; DB 9; Length 211;
 Best Local Similarity 100.0%; Pred. No. 5e-09;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIHLLTSLAGTCSAHVQQLA 23
 DB 3 KIHLLTSLAGTCSAHVQQLA 25

RESULT 4
 ADB16910
 ID ADB16910 standard; peptide; 22 AA.
 XX
 AC ADB16910;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 XX Escherichia coli YCPC penetrating peptide 32.
 DE
 XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 XX infertility; hormone; vitamin deficiency; neurodegenerative;
 KM cardiovascular; hematological; endocrine disorder; obesity;
 KM neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
 KM osteopathic; cytostatic; nootropic.
 XX

OS Escherichia coli.
 XX
 XX WO2003066859-A2.
 XX
 XX 14-AUG-2003.
 PD
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 PP
 XX
 PR 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA
 XX Ben-Sasson SA, Cohen E;
 PI WPI, 2003-697452/66.
 DR
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PT
 XX
 XX Claim 2; Page 15; 60pp; English.
 PS
 XX
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, hematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 CC cytostatic or nootropic activities. This peptide is from YCPC of
 CC Escherichia coli and is penetrating peptide 32 of the invention.
 XX
 XX Sequence 22 AA;
 SQ
 Query Match 71.4%; Score 80; DB 6; Length 22;
 Best Local Similarity 89.5%; Pred. No. 3.5e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITTLALGTCQSARLVQQLA 23
 DB 4 ITTLALGTCQSARLVQQLA 22

RESULT 5
 AEB08244
 ID AEB08244 standard; peptide; 22 AA.
 XX
 AC AEB08244;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 XX Escherichia coli YCPC penetrating peptide 32. SEQ ID NO: 27.
 DE
 XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 KM metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KM degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KM neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KM nootropic; neurological disease; parkinsons disease; antiparkinsonian;
 KM dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KM cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KM coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 KM nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KM

KM genital disease; hematological disease; antianemic; anemia;
 KM autoimmune disease; immunosuppressive; immune deficiency;
 KM immunostimulant; infectious disease; antimicrobial; infection;
 KM erectile dysfunction; andrology; major depressive disorder;
 KM antidepressant; psychiatric disorder; pain; analgesic;
 KM bacterial infection; antibacterial; viral infection; virucide;
 KM fungal infection; fungicide; parasitic infection; antiparasitic;
 KM renal failure; antifertility; antirheumatic; cytostatic;
 KM antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX Becherichia coli.
 OS
 PN US2005136103-A1.
 XX
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 PI
 DR WPI; 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprising
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 27; 59pp; English.

CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present invention is the Becherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 CC
 XX
 SQ Sequence 22 AA;
 Query Match 71.4%; Score 80; DB 9; Length 22;
 Best Local Similarity 89.5%; Pred. No. 3.5e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITLSIAGICQSAHLVQOLA 23
 |||:|||||
 Db 4 ITLALAGICQSAHLVQOLA 22

RESULT 6
 ADB16888

ID ADB16888 standard; peptide; 23 AA.
 XX
 AC ADB16888;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Becherichia coli YCFC penetrating peptide 3.
 XX
 KM penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 KM infertility; hormone; vitamin deficiency; neurodegenerative;
 KM cardiovascular; haematological; endocrine disorder; obesity;
 KM neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
 KM osteopathic; cyostatic; nootropic.
 XX
 OS Becherichia coli.
 XX
 PN W02003066859-A2.
 XX
 PD 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 PR
 PR 07-FEB-2002; 2002US-0355396P.
 XX
 PR (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 PA Ben-Sasson SA, Cohen E;
 PI
 PI Ben-Sasson SA, Cohen E;
 PI
 DR WPI; 2003-697452/66.
 XX
 PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 PS Claim 2; Page 14; 60pp; English.

CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 CC cyostatic or nootropic activities. This peptide is from YCFC of
 CC Becherichia coli and is penetrating peptide 3 of the invention.
 CC
 XX
 SQ Sequence 23 AA;
 Query Match 71.4%; Score 80; DB 6; Length 23;
 Best Local Similarity 89.5%; Pred. No. 3.7e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITLSIAGICQSAHLVQOLA 23
 |||:|||||
 Db 5 ITLALAGICQSAHLVQOLA 23

RESULT 7
 ADB16909
 ID ADB16909 standard; peptide; 23 AA.
 AC ADB16909;
 XX
 DT 20-NOV-2003 (first entry)

XX Escherichia coli YCFC penetrating peptide 31.
 XX
 XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 XX intercellular; hormone; vitamin deficiency; neurodegenerative;
 XX cardiovascular; haematological; endocrine disorder; obesity;
 XX neoplastic disease; neuroprotective; cardiac; antidiabetic; osteoporotic;
 XX osteopathic; cytoskeletal; neurotropic.
 XX
 XX Escherichia coli.
 XX
 XX WO200306659-A2.
 XX
 XX 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-1B000966.
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX
 XX WPI; 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX
 XX Claim 2, Page 15, 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, haematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX Accordingly, the peptides of this invention can be used in compositions
 XX that have neuroprotective, cardiac, antidiabetic, osteoporotic,
 XX cytoskeletal or neurotropic activities. This peptide is from YCFC of
 XX Escherichia coli and is penetrating peptide 31 of the invention.
 XX
 XX Sequence 23 AA;
 XX
 XX Query Match 71.4%; Score 80; DB 6; Length 23;
 XX Best Local Similarity 89.5%; Pred. No. 3.7e-05;
 XX Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 XX
 XX 5 ITTSLAGIGOSARLVQOLA 23
 XX |||||
 XX 5 ITTSLAGIGOSARLVQOLA 23
 XX
 XX RESULT 8
 XX AEB08220 ID AEB08220 standard; peptide; 23 AA.
 XX
 XX AEB08220;
 XX
 XX 25-AUG-2005 (first entry)
 XX
 XX Escherichia coli YCFC penetrating peptide 3, SEQ ID NO: 3.
 XX
 XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX

XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antidiabetic; osteoporotic;
 XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephropathic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; virucide;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antiinfectivity; antineumatic; cytoskeletal;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX Escherichia coli.
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX
 XX 17-SEP-2003; 2003US-00665184.
 XX
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHR/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX
 XX WPI; 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 53; SEQ ID NO 3; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for transcellular delivery of effector. The
 XX invention comprises the effector sequentially coupled with a counter ion
 XX and at least one hydrophobic agent, where the effector is selectively
 XX encapsulated into a complex. The invention is useful for translocating
 XX effectors across a biological barrier such as epithelial cells and
 XX endothelial cells during treatment and prevention of disease or
 XX pathological conditions (including endocrine disorders, diabetes,
 XX infertility, hormone deficiencies, osteoporosis, ophthalmological
 XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 XX Parkinson's disease, multiple sclerosis, Huntington's disease,
 XX cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 XX coagulable states, coronary disease, cerebrovascular events, metabolic
 XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 XX hematological disorders, anemia of different entities, immunologic and
 XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
 XX infectious diseases, viral infections, bacterial infections, fungal
 XX infections, parasitic infections, neoplastic diseases, multi-factorial
 XX disorders, impotence, chronic pain, depression, different fibrosis states
 XX and short stature) and for mucosal vaccination against anthrax and
 XX hepatitis B. The present sequence is the Escherichia coli YCFC
 XX penetrating peptide. This sequence is used in the effective translocation
 XX of aminoacyl-coester antibiotics and antifungal agents across an epithelial
 XX barrier.
 XX
 XX Sequence 23 AA;
 XX
 XX Query Match 71.4%; Score 80; DB 9; Length 23;
 XX Best Local Similarity 89.5%; Pred. No. 3.7e-05;
 XX

Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ITTSLAGICGSAHVVQOLA 23
 ||:|||||
 Db 5 ITTALAGICGSAHVVQOLA 23

RESULT 9
 AEB08243 standard; peptide; 23 AA.

AC AEB08243;
 XX
 DT 25-AUG-2005 (first entry)
 XX

DB Escherichia coli YCFC penetrating peptide 31, SEQ ID NO: 26.

XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathol;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.

XX OS Escherichia coli.
 XX US2005136103-A1.
 XX 23-JUN-2005.
 XX PD 16-SEP-2004; 2004US-00942300.
 XX PF 17-SEP-2003; 2003US-00664989.
 XX PR 17-SEP-2003; 2003US-00665184.
 XX PR 17-SEP-2003; 2003US-0503615P.
 XX PA (BENS/) BEN-SASSON S A.
 XX (COHE/) COHEN E.
 XX PI Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 26; 59pp; English.

CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for a transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,

CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematologic disorders, autoimmune diseases, immune deficiencies, and
 CC rheumatologic disorders, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.

XX
 SQ Sequence 23 AA;

Query Match 71.4%; Score 80; DB 9; Length 23;
 Best Local Similarity 89.5%; Pred. No. 3.7e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ITTSLAGICGSAHVVQOLA 23
 ||:|||||
 Db 5 ITTALAGICGSAHVVQOLA 23

RESULT 10
 ADB16908
 ID ADB16908 standard; peptide; 24 AA.
 XX
 AC ADB16908;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Escherichia coli YCFC penetrating peptide 30.

XX
 KW penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 KW infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathol; cytostatic; neurotropic.

XX OS Escherichia coli.
 XX WO2003066859-A2.
 XX 14-AUG-2003.
 XX PD 07-FEB-2003; 2003WO-1B000968.
 XX PF 07-FEB-2003; 2003US-0355396P.
 XX PR 07-FEB-2002; 2002US-0355396P.
 XX PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX PI Ben-Sasson SA, Cohen E;
 XX WPI; 2003-697452/66.
 XX
 PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 PS Claim 2; Page 15; 60pp; English.

CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, hematological and endocrine disorders, as well as obesity and neoplastic disease. Accordingly, the peptides of this invention can be used in compositions that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic, cyostatic or neurotropic activities. This peptide is from YCPC of Escherichia coli and is penetrating peptide 30 of the invention.

Sequence 24 AA:

Query Match 71.4%; Score 80; DB 6; Length 24;
Best Local Similarity 89.5%; Pred. No. 3.9e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITLSLNGICQSAHLVQQLA 23
Db 6 ITLSLNGICQSAHLVQQLA 24

RESULT 11

AE808242 standard; peptide, 24 AA.

AC AE808242;

DT 25-AUG-2005 (first entry)

DE Escherichia coli YCPC penetrating peptide 30, SEQ ID NO: 25.

pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes; antidiabetic; endocrine disease; gastrointestinal disease; metabolic disorder; hormone deficiency; osteoporosis; osteopathic; degenerative; musculoskeletal disease; ocular disease; ophthalmological; neurodegenerative disease; neuroprotective; Alzheimer's disease; neurotropic; neurological disease; Parkinson's disease; antiparkinsonian; dementia; multiple sclerosis; immune disorder; Huntingtons chorea; anticonvulsant; genetic disorder; cardiovascular disease; cardiovascular-gen.; atherosclerosis; antiarteriosclerotic; coronary artery disease; cardiac; vasotropic; obesity; anorectic; nutritional disorder; vitamin deficiency; renal disease; nephrotropic; genitourinary disease; hematological disease; antiemetic; anemia; autoimmune disease; immunosuppressive; immune deficiency; immunosuppressant; infectious disease; antimicrobial; infection; erectile dysfunction; andrology; major depressive disorder; antidepressant; psychiatric disorder; pain; analgesic; bacterial infection; antibacterial; viral infection; virucide; fungal infection; fungicide; parasitic infection; antiparasitic; renal failure; antifertility; antipneumatic; cytostatic; antiinflammatory; hepatotropic; hepatitis B virus infection.

Escherichia coli.

US2005136103-A1.

23-JUN-2005.

16-SEP-2004; 2004US-00942300.

17-SEP-2003; 2003US-00664989.

17-SEP-2003; 2003US-00665184.

17-SEP-2003; 2003US-0503615P.

(BENS/) BEN-SASSON S. A.

(COHE/) COHEN E.

Ben-Sasson SA, Cohen E;

WPI, 2005-444089/45.

Composition used for translocating effectors across barrier such as epithelial cells during treatment of e.g. endocrine disorders comprises

effector sequentially coupled with counter ion and hydrophobic agent.

Claim 46; SEQ ID NO 25; 59pp; English.

The present invention relates to a pharmaceutical composition of penetrating peptides for transmembrane delivery of effector. The invention comprises the effector sequentially coupled with a counter ion and at least one hydrophobic agent, where the effector is selectively encapsulated into a complex. The invention is useful for translocating effectors across a biological barrier such as epithelial cells and endothelial cells during treatment and prevention of disease or pathological conditions (including endocrine disorders, diabetes, infertility, hormone deficiencies, osteoporosis, ophthalmological disorders, neurodegenerative disorders, Alzheimer's disease, dementia, Parkinson's disease, multiple sclerosis, Huntington's disease, hypocoagulable states, coronary disease, cerebrovascular events, metabolic disorders, obesity, vitamin deficiencies, renal disorders, renal failure, hematological disorders, anemia of different entities, immunologic and rheumatologic disorders, autoimmune diseases, immune deficiencies, infectious diseases, viral infections, bacterial infections, fungal infections, parasitic infections, neoplastic diseases, multi-factorial disorders, impotence, chronic pain, depression, different fibrosis states and short stature) and for mucosal vaccination against anthrax and hepatitis B. The present sequence is the Escherichia coli YCPC penetrating peptide. This sequence is used in the effective translocation of aminoglycoside antibiotics and antifungal agents across an epithelial barrier.

Sequence 24 AA:

Query Match 71.4%; Score 80; DB 9; Length 24;
Best Local Similarity 89.5%; Pred. No. 3.9e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITLSLNGICQSAHLVQQLA 23
Db 6 ITLSLNGICQSAHLVQQLA 24

RESULT 12

ADB16922 standard; peptide, 25 AA.

AC ADB16922;

DT 20-NOV-2003 (first entry)

DE Escherichia coli peptide 3 linked to recombinant human insulin.

recombinant human insulin; epithelial; endothelial; tight junction; diabetes; infertility; hormone; vitamin deficiency; neurodegenerative; cardiovascular; hematological; endocrine disorder; obesity; neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic; osteopathic; cytostatic; neurotropic; penetrating peptide; diabetes.

Synthetic.

Escherichia coli.

Key Location/Qualifiers

Peptide 1..23 "Penetrating peptide 3"

Modified-site 25 /note="Recombinant human insulin peptide coupled to the C-terminus"

MO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-IB000968.

07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA Ben-Sasson SA, Cohen E;
 PI WPI; 2003-697452/66.
 DR
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 PS Example 8; Page 42; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antidiabetic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide sequence consists of the
 CC Escherichia coli penetrating peptide 3 coupled to recombinant human
 CC insulin in the absence of a detachable linker peptide as a cleavage site,
 CC used to deliver insulin across mucosal epithelia as a treatment for
 CC diabetes, in an exemplification of the invention.
 XX
 SQ Sequence 25 AA;
 Query Match 71.4%; Score 80; DB 6; Length 25;
 Best Local Similarity 89.5%; Pred. No. 4.1e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 5 ITTSLAGICQSAHLVQOLA 23
 DB 5 ITTALAGICQSAHLVQOLA 23
 RESULT 13
 ADB16906
 ID ADB16906 standard; peptide; 25 AA.
 AC
 XX ADB16906;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Escherichia coli peptide 3 linked to linearised insulin receptor peptide.
 XX
 KW linearised insulin receptor; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiant; antidiabetic/osteopathic;
 KW osteopathic; cyostatic; neurotropic; penetrating peptide; diabetes.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Modified-site /note= "Penetrating peptide 3"
 FT 25
 FT /note= "Linearised insulin receptor peptide coupled to
 FT the C-terminus"
 XX
 PN WO2003066859-A2.

XX 14-AUG-2003.
 XX
 PD 07-FEB-2003; 2003MO-IB000968.
 XX
 PF 07-FEB-2002; 2002US-0355396P.
 XX
 PR (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA Ben-Sasson SA, Cohen E;
 PI WPI; 2003-697452/66.
 DR
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 PS Example 9; Page 43; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antidiabetic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide sequence consists of the
 CC Escherichia coli penetrating peptide 3 coupled to a linearised insulin
 CC receptor, used to deliver insulin across mucosal epithelia as a treatment
 CC for diabetes, in an exemplification of the invention.
 XX
 SQ Sequence 25 AA;
 Query Match 71.4%; Score 80; DB 6; Length 25;
 Best Local Similarity 89.5%; Pred. No. 4.1e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 5 ITTSLAGICQSAHLVQOLA 23
 DB 5 ITTALAGICQSAHLVQOLA 23
 RESULT 14
 ADB16904
 ID ADB16904 standard; peptide; 26 AA.
 AC
 XX ADB16904;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE E.coli penetrating peptide 3 linked to heparin without a cleavage site.
 XX
 KW heparin; epithelial; endothelial; tight junction; diabetes; infertility;
 KW hormone; vitamin deficiency; neurodegenerative; cardiovascular;
 KW haematological; endocrine disorder; obesity; neoplastic disease;
 KW neuroprotective; cardiant; antidiabetic/osteopathic; cyostatic;
 KW neurotropic; penetrating peptide.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "Penetrating peptide 3"

FT Modified-site 26 /note= "Heparin peptide coupled to the C-terminus"
 FT
 XX
 XX WO2003066859-A2.
 PN
 XX 14-AUG-2003.
 PD
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 PP
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 PR
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA
 XX Ben-Sasson SA, Cohen E;
 PI
 XX MPI, 2003-697452/66.
 DR
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PT
 XX
 XX Example 10; Page 44; 60pp; English.
 PS
 XX
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticosugulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, hematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antidiabetic, osteopathic,
 CC cyostatic or neurotrophic activities. This peptide sequence consists of the
 CC Bacterichia coli penetrating peptide 3 coupled to heparin in the absence
 CC of a detachable linker peptide as a cleavage site, used to deliver low
 CC molecular weight heparin across mucosal epithelia in an exemplification
 CC of the invention.
 CC
 XX Sequence 26 AA;
 SQ
 Query Match 71.4%; Score 80; DB 6; Length 26;
 Best Local Similarity 89.5%; Pred. No. 4.3e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 5 ITTSLAGTGSARLVQOLA 23
 DB 5 ITTALAGTGSARLVQOLA 23
 RESULT 15
 ADB16920
 ID ADB16920 standard; peptide; 29 AA.
 AC ADB16920;
 XX
 XX 20-NOV-2003 (first entry)
 DT
 XX
 XX Escherichia coli peptide 3 coupled to imaging compound linker, IBM-002V3.
 DE
 XX epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiac; antidiabetic; osteopathic; cyostatic; neurotrophic;
 KW imaging linker; penetrating peptide; IBM-002V3.
 XX
 OS Synthetic.

OS Escherichia coli.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 29 /note= "C-terminal amide"
 FT
 XX
 XX WO2003066859-A2.
 PN
 XX 14-AUG-2003.
 PD
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 PP
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 PR
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA
 XX Ben-Sasson SA, Cohen E;
 PI
 XX MPI, 2003-697452/66.
 DR
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PT
 XX
 XX Example 3; Page 40; 60pp; English.
 PS
 XX
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticosugulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, hematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antidiabetic, osteopathic,
 CC cyostatic or neurotrophic activities. This peptide sequence is IBM-002V3,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 CC
 XX Sequence 29 AA;
 SQ
 Query Match 71.4%; Score 80; DB 6; Length 29;
 Best Local Similarity 89.5%; Pred. No. 4.8e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 5 ITTSLAGTGSARLVQOLA 23
 DB 4 ITTALAGTGSARLVQOLA 22
 Search completed: January 23, 2006, 09:20:47
 Job time : 69.2 secs

This Page Blank (jsp:do)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 Seconds
(without alignments)
114.551 Million cell updates/sec

Title: US-10-501-838a-5
Sequence: 1 KIHLLTSLAGICOSAHLVQOLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*
1: /cgn2_6/prodata/1/1aa/5-COMB.pep:*
2: /cgn2_6/prodata/1/1aa/6-COMB.pep:*
3: /cgn2_6/prodata/1/1aa/H-COMB.pep:*
4: /cgn2_6/prodata/1/1aa/BCUS-COMB.pep:*
5: /cgn2_6/prodata/1/1aa/RE-COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillset1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	67.9	217	2	US-09-489-039A-7789 Sequence 7789, Ap
2	72	64.3	216	2	US-09-543-681A-7993 Sequence 7993, Ap
3	44	39.3	210	2	US-09-252-991A-21085 Sequence 21085, A
4	43	38.4	64	2	US-09-149-476-323 Sequence 323, App
5	43	38.4	81	2	US-09-513-999C-4118 Sequence 4118, Ap
6	42.5	37.9	234	2	US-09-489-039A-9787 Sequence 9787, Ap
7	42	37.5	219	2	US-09-543-681A-6092 Sequence 6092, Ap
8	42	37.5	426	2	US-09-268-311-2 Sequence 2, Appl1
9	42	37.5	426	2	US-09-268-311-3 Sequence 2, Appl1
10	42	37.5	426	2	US-09-154-219-2 Sequence 2, Appl1
11	42	37.5	426	2	US-09-154-219-3 Sequence 3, Appl1
12	42	37.5	426	2	US-09-949-016-6936 Sequence 6936, Ap
13	42	37.5	426	2	US-09-796-844-2 Sequence 2, Appl1
14	42	37.5	426	2	US-09-796-844-3 Sequence 3, Appl1
15	42	37.5	433	2	US-09-268-311-18 Sequence 18, Appl
16	42	37.5	433	2	US-09-796-844-18 Sequence 18, Appl
17	42	37.5	457	2	US-09-252-991A-19217 Sequence 19217, A
18	42	37.5	502	2	US-09-747-259-12 Sequence 12, Appl
19	42	37.5	502	2	US-09-816-744-12 Sequence 12, Appl
20	42	37.5	504	2	US-09-949-016-11658 Sequence 11658, A
21	42	37.5	536	2	US-09-902-540-10280 Sequence 10280, A
22	42	37.5	567	2	US-09-538-092-1365 Sequence 1365, Ap
23	42	37.5	604	2	US-09-134-000C-5828 Sequence 5828, Ap
24	41	36.6	430	1	US-08-846-762-73 Sequence 73, Appl
25	41	36.6	1336	2	US-09-949-016-9879 Sequence 9879, Ap
26	41	36.6	1695	2	US-09-866-108A-15753 Sequence 15753, A
27	40	35.7	51	2	US-09-513-999C-7066 Sequence 7066, Ap

28	40	35.7	253	2	US-09-602-787A-236 Sequence 236, App
29	40	35.7	284	2	US-09-265-585C-111 Sequence 111, App
30	40	35.7	284	2	US-09-265-585C-127 Sequence 127, App
31	40	35.7	285	2	US-09-186-276B-54 Sequence 54, Appl
32	40	35.7	285	2	US-08-842-445-54 Sequence 54, Appl
33	40	35.7	285	2	US-09-186-188B-54 Sequence 54, Appl
34	40	35.7	285	2	US-09-265-585C-54 Sequence 54, Appl
35	40	35.7	453	2	US-09-252-991A-32168 Sequence 32168, A
36	40	35.7	823	2	US-09-248-796A-16699 Sequence 16699, A
37	40	35.7	3421	2	US-09-452-638-53 Sequence 53, Appl
38	40	35.7	3421	2	US-09-121-587A-13 Sequence 13, Appl
39	39	34.8	29	2	US-09-270-767-59495 Sequence 59495, A
40	39	34.8	150	2	US-09-252-991A-20686 Sequence 20686, A
41	39	34.8	189	2	US-09-270-767-44080 Sequence 44080, A
42	39	34.8	192	1	US-08-086-428B-102 Sequence 102, App
43	39	34.8	192	1	US-08-468-192-102 Sequence 102, App
44	39	34.8	192	1	US-08-290-665A-102 Sequence 102, App
45	39	34.8	192	2	US-08-466-601A-102 Sequence 102, App

ALIGNMENTS

```
RESULT 1
US-09-489-039A-7789
; Sequence 7789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7789
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match      67.9%; Score 76; DB 2; Length 217;
Best Local Similarity 78.9%; Pred. No. 0.0007;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      5 ITLTLGVCQARLVQOLA 23
      |||:||||:|||||
DB      12 ITLTLGVCQARLVQOLA 30

RESULT 2
US-09-543-681A-7993
; Sequence 7993, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543, 681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7993
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7993

Query Match      64.3%; Score 72; DB 2; Length 216;
```

Best Local Similarity 73.7%; Pred. No. 0.0028;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITTSLAGICQSAHVVQQLA 23
Db 15 ITTSLAGICQSAHVVQQLA 33

RESULT 3
US-09-252-991A-21085
Sequence 21085, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21085
LENGTH: 210
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21085

Query Match 39.3%; Score 44; DB 2; Length 210;
Best Local Similarity 40.0%; Pred. No. 50;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 3 HLTSLAGICQSAHVVQQL 22
Db 114 HVLKRSYGCLCTAVAVTEL 133

RESULT 4
US-09-149-476-323
Sequence 323, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878

```

/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,662
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,872
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,882
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,637
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,903
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,888
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,879
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,880
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,894
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,911
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,636
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,874
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,910
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,864
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,631
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,845
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,892
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/057,761
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/047,595
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,599
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,588
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,585
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,590
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,594
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,589
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,593
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/047,614
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/043,578
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/043,576
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/047,501
/ EARLIER FILING DATE: 1997-05-23
/ EARLIER APPLICATION NUMBER: 60/043,670
/ EARLIER FILING DATE: 1997-04-11
/ EARLIER APPLICATION NUMBER: 60/056,632
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,664
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,876
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,881
/ EARLIER FILING DATE: 1997-08-22

```

```

/ EARLIER APPLICATION NUMBER: 60/056,909
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,875
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,862
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,887
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/056,908
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/048,964
/ EARLIER FILING DATE: 1997-06-06
/ EARLIER APPLICATION NUMBER: 60/057,650
/ EARLIER FILING DATE: 1997-09-05
/ EARLIER APPLICATION NUMBER: 60/056,884
/ EARLIER FILING DATE: 1997-08-22
/ EARLIER APPLICATION NUMBER: 60/057,669
/ EARLIER FILING DATE: 1997-09-05
/ EARLIER APPLICATION NUMBER: 60/049,610
/ EARLIER FILING DATE: 1997-06-13
/ EARLIER APPLICATION NUMBER: 60/061,060
/ EARLIER FILING DATE: 1997-10-02

```

```

Query Match          38.4% Score 43; DB 2; Length 64;
Best Local Similarity 46.7% Pred. No. 19;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

Qy      3 HLITSLAGICQSAH 17
||:||||:|
Db      25 HLITSLAGICQYH 39

```

```

RESULT 5
US-09-513-999C-4118
Sequence 4118, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.
APPLICANT: Duciart, A.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961
FILE REFERENCE: 59, US2, REG
CURRENT APPLICATION NUMBER: US/09/513,999C
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 4118
LENGTH: 81
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -19..-1
OTHER INFORMATION: score 13.3
OTHER INFORMATION: seq LALLAVLQGVCA/EV
FEATURE:
NAME/KEY: UNSURE
LOCATION: -16
OTHER INFORMATION: Xaa=Ala or Thr
US-09-513-999C-4118

```

```

Query Match          38.4% Score 43; DB 2; Length 81;
Best Local Similarity 47.1% Pred. No. 24;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
Qy      4 LITSLAGICQSAH 20
||:||||:|
Db      9 LITSLAGICQYH 25

```

```

Query Match Similarity 37.5% Score 42; DB 2; length 219;
Best Local Similarity 36.8%; Pred. No. 1e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0

QY          3 HILITSLAGICQSAHLVQ 21
      | | | | | | | | | |
Db          54 HISSDLGRTCTATRIAQ 72

RESULT 8
US-09-268-311-2
; Sequence 2, Application US/09268311
; Patent No. 6482923
; GENERAL INFORMATION:
; APPLICANT: Shi, Yangu
; APPLICANT: Ruben, Steve M.
; TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
; FILE REFERENCE: PF398Pl
; CURRENT APPLICATION NUMBER: US/09/268,311
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 60/059,133
; EARLIER FILING DATE: 1997-09-17
; EARLIER APPLICATION NUMBER: 09/154,219
; EARLIER FILING DATE: 1998-09-16

```

```

RESULT 10
US-09-154-219-2
; Sequence 2, Application US/09154219
; Patent No. 6635443
; GENERAL INFORMATION:
; APPLICANT: Shi, Yanguu
; APPLICANT: Ruben, Steve M.
; TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
; FILE REFERENCE: pp398
; CURRENT APPLICATION NUMBER: US/09/154,219
; CURRENT FILING DATE: 1998-09-16
; EARLIER APPLICATION NUMBER: 60/059,133
; EARLIER FILING DATE: 1997-09-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-154-219-2

Query March          37.5%   Score 42;  DB 2;  Length 426;
Best Local Similarity 61.5%   Pred. No. 2.2e+02;
Matches      8;  Conservative      3;  Mismatches      2;  Indels      0;  Gaps      0

```

OY 4 LITSLAGICQSA 16
|:||||:|:
DB 3 LVLSLALCRSA 15

RESULT 11

US-09-154-219-3
Sequence 3, Application US/09154219
Patent No. 6635443
GENERAL INFORMATION:
APPLICANT: Shi, Yangu
APPLICANT: Ruben, Steve M.
TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
FILE REFERENCE: P398
CURRENT APPLICATION NUMBER: US/09/154,219
CURRENT FILING DATE: 1998-09-16
EARLIER APPLICATION NUMBER: 60/059,133
PRIOR FILING DATE: 1997-09-17
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
US-09-154-219-3

Query Match 37.5%; Score 42; DB 2; Length 426;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 LITSLAGICQSA 16
|:||||:|:
DB 3 LVLSLALCRSA 15

RESULT 12

US-09-949-016-6936
Sequence 6936, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6936
LENGTH: 426
TYPE: PRT
ORGANISM: Human
US-09-949-016-6936

Query Match 37.5%; Score 42; DB 2; Length 426;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 LITSLAGICQSA 16
|:||||:|:
DB 3 LVLSLALCRSA 15

RESULT 13
US-09-796-844-2
Sequence 2, Application US/09796844

Patent No. 6649719

GENERAL INFORMATION:
APPLICANT: Shi, Yangu
APPLICANT: Ruben, Steve M.
TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
FILE REFERENCE: P398P2
CURRENT APPLICATION NUMBER: US/09/796,844
CURRENT FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/187,015
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: PCT/US00/05759
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: PCT/US99/21048
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: 09/268,311
PRIOR FILING DATE: 1999-03-16
PRIOR APPLICATION NUMBER: PCT/US98/19121
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 09/154,219
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/059,133
PRIOR FILING DATE: 1997-09-17
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-844-2

Query Match 37.5%; Score 42; DB 2; Length 426;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 4 LITSLAGICQSA 16
|:||||:|:
DB 3 LVLSLALCRSA 15

RESULT 14

US-09-796-844-3
Sequence 3, Application US/09796844
Patent No. 6649719
GENERAL INFORMATION:
APPLICANT: Shi, Yangu
APPLICANT: Ruben, Steve M.
TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
FILE REFERENCE: P398P2
CURRENT APPLICATION NUMBER: US/09/796,844
CURRENT FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/187,015
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: PCT/US00/05759
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: PCT/US99/21048
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: 09/268,311
PRIOR FILING DATE: 1999-03-16
PRIOR APPLICATION NUMBER: PCT/US98/19121
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 09/154,219
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/059,133
PRIOR FILING DATE: 1997-09-17
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 426
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-844-3

Query Match 37.5%; Score 42; DB 2; Length 426;

Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LITSLAGICQSA 16
|:||||:|
Db 3 LVLSTLALCRSA 15

RESULT 15
US-09-268-311-18
; Sequence 18, Application US/09268311
; Patent No. 6482923
; GENERAL INFORMATION:
; APPLICANT: Ruben, Steve M.
; TITLE OF INVENTION: Interleukin 17 Receptor-Like Protein
; FILE REFERENCE: PF398P1
; CURRENT APPLICATION NUMBER: US/09/268,311
; EARLIER FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 60/059,133
; EARLIER FILING DATE: 1997-09-17
; EARLIER APPLICATION NUMBER: 09/154,219
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-268-311-18

Query Match 37.5%; Score 42; DB 2; Length 433;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LITSLAGICQSA 16
|:||||:|
Db 3 LVLSTLALCRSA 15

Search completed: January 23, 2006, 09:35:46
Job time : 17.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:56 Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838a-5
Perfect score: 112
Sequence: 1 KIHLLTSLAGICQSAHLVQOLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues
Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA_New.*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42.5	37.9	365	7 US-11-000-463-241	Sequence 241, App
2	42.5	37.9	365	7 US-11-000-463-713	Sequence 713, App
3	42	37.5	502	6 US-10-063-703-158	Sequence 158, App
4	42	37.5	502	7 US-11-102-240-158	Sequence 158, App
5	41	36.6	293	6 US-10-967-671-15	Sequence 15, App
6	40	35.7	291	6 US-10-467-657-644	Sequence 644, App
7	39	34.8	42	6 US-10-512-295A-3	Sequence 3, App
8	39	34.8	687	6 US-10-055-877-203	Sequence 203, App
9	38	33.9	29	7 US-11-079-139-12	Sequence 12, App
10	38	33.9	307	7 US-11-019-711-87	Sequence 87, App
11	38	33.9	307	7 US-11-019-711-88	Sequence 88, App
12	37.5	33.5	256	7 US-11-179-977-15	Sequence 15, App
13	37	33.0	32	7 US-11-033-039-1145	Sequence 1145, App
14	37	33.0	143	6 US-10-793-626-370	Sequence 370, App
15	37	33.0	462	6 US-10-793-626-906	Sequence 906, App
16	37	33.0	655	6 US-10-055-877-40	Sequence 40, App
17	37	33.0	668	6 US-10-997-201A-30	Sequence 30, App
18	37	33.0	678	6 US-10-055-877-202	Sequence 202, App
19	37	33.0	687	6 US-10-055-877-199	Sequence 199, App
20	37	33.0	687	6 US-10-055-877-201	Sequence 201, App
21	37	33.0	1255	7 US-11-022-562-235	Sequence 235, App
22	37	33.0	1255	7 US-11-052-554A-265	Sequence 265, App
23	37	33.0	1255	7 US-11-052-554A-266	Sequence 266, App
24	37	33.0	1255	7 US-11-052-554A-267	Sequence 267, App
25	37	33.0	1255	7 US-11-052-554A-268	Sequence 268, App

26	37	33.0	1255	7 US-11-052-554A-269	Sequence 269, App
27	37	33.0	1255	7 US-11-052-554A-270	Sequence 270, App
28	37	33.0	1255	7 US-11-052-554A-271	Sequence 271, App
29	37	33.0	1255	7 US-11-052-554A-272	Sequence 272, App
30	37	33.0	1255	7 US-11-052-554A-273	Sequence 273, App
31	37	33.0	1255	7 US-11-052-554A-274	Sequence 274, App
32	37	33.0	1255	7 US-11-052-554A-275	Sequence 275, App
33	37	33.0	1279	6 US-10-957-880-3	Sequence 3, App
34	36.5	32.6	168	6 US-10-467-657-2154	Sequence 2154, App
35	36	32.1	24	6 US-10-986-501-122	Sequence 122, App
36	36	32.1	31	7 US-11-021-441-51	Sequence 51, App
37	36	32.1	96	5 US-09-978-360A-654	Sequence 654, App
38	36	32.1	143	7 US-11-102-883-12	Sequence 12, App
39	36	32.1	235	6 US-10-821-234-1111	Sequence 1111, App
40	36	32.1	263	7 US-11-102-883-36	Sequence 36, App
41	36	32.1	274	7 US-11-102-883-34	Sequence 34, App
42	36	32.1	283	7 US-11-102-883-35	Sequence 35, App
43	36	32.1	288	7 US-11-052-554A-353	Sequence 2, App
44	36	32.1	289	7 US-11-102-883-2	Sequence 26, App
45	36	32.1	289	7 US-11-102-883-26	Sequence 26, App

ALIGNMENTS

RESULT 1
US-11-000-463-241
Sequence 241, Application US/11000463
Publication No. US20050266423A1
GENERAL INFORMATION:
APPLICANT: Tang, Y Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Cao, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIPACN
CURRENT APPLICATION NUMBER: US/11/000,463
CURRENT FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 241
LENGTH: 365
TYPE: PRT
ORGANISM: Homo sapiens
US-11-000-463-241
Query Match 37.9%; Score 42.5; DB 7; Length 365;
Best Local Similarity 45.5%; Pred. No. 9;
Matches 10; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

Db 68 LMLGLAGFCGSMASGHLFKQMA 89

RESULT 2

US-11-000-463-713
; Sequence 713, Application US/11000463
; Publication No. US20050266423A1

GENERAL INFORMATION:

APPLICANT: Tang, Y Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Chen, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIP4CN

CURRENT APPLICATION NUMBER: US/11/000,463

PRIOR FILING DATE: 2004-11-29

PRIOR APPLICATION NUMBER: 10/291,265

PRIOR FILING DATE: 2002-11-08

PRIOR APPLICATION NUMBER: PCT/US01/02623

PRIOR FILING DATE: 2001-01-25

PRIOR APPLICATION NUMBER: 09/922,279

PRIOR FILING DATE: 2001-08-03

PRIOR APPLICATION NUMBER: 09/491,404

PRIOR FILING DATE: 2000-01-25

PRIOR APPLICATION NUMBER: 09/617,746

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 09/631,451

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 09/633,870

PRIOR FILING DATE: 2000-09-15

NUMBER OF SEQ ID NOS: 944

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 713

LENGTH: 365

TYPE: PRT

ORGANISM: Homo sapiens

US-11-000-463-713

Query Match 37.9%; Score 42.5; DB 7; Length 365;

Best Local Similarity 45.5%; Pred. No. 9;

Matches 10; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

Qy 5 LITSLAGIC--QSAAHYQOLA 23

Db 68 LMLGLAGFCGSMASGHLFKQMA 89

RESULT 3

US-10-063-703-158

; Sequence 158, Application US/10063703

; Publication No. US20060008901A1

GENERAL INFORMATION:

APPLICANT: Baton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerltzen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,703

; Prior Application removed - See Palm or File Wrapper

NUMBER OF SEQ ID NOS: 170

SEQ ID NO 158

LENGTH: 502

TYPE: PRT

ORGANISM: Homo Sapien

US-10-063-703-158

Query Match 37.5%; Score 42; DB 6; Length 502;

Best Local Similarity 61.5%; Pred. No. 15;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LITSLAGICQSA 16

Db 3 LVTLSTAAACRSA 15

RESULT 4

US-11-102-240-158

; Sequence 158, Application US/11102240

; Publication No. US20050260647A1

GENERAL INFORMATION:

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
FILE REFERENCE: P3230R1C106C
CURRENT APPLICATION NUMBER: US/11/102,240

PRIOR FILING DATE: 2005-04-08

PRIOR APPLICATION NUMBER: 10/063662

PRIOR FILING DATE: 2002-05-07

PRIOR APPLICATION NUMBER: 10/006867

PRIOR FILING DATE: 2001-12-06

PRIOR APPLICATION NUMBER: PCT/US00/23328

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: 60/170262

PRIOR FILING DATE: 199-12-09

NUMBER OF SEQ ID NOS: 170

SEQ ID NO 158

LENGTH: 502

TYPE: PRT

ORGANISM: Homo Sapien

US-11-102-240-158

Query Match 37.5%; Score 42; DB 7; Length 502;

Best Local Similarity 61.5%; Pred. No. 15;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LITSLAGICQSA 16

Db 3 LVTLSTAAACRSA 15

RESULT 5

US-10-967-671-15

; Sequence 15, Application US/10967671

; Publication No. US20050245728A1

GENERAL INFORMATION:

APPLICANT: EDWARDS, ALDI
APPLICANT: DHARMSI, AKIL
APPLICANT: VERDADY, MASOUD
APPLICANT: KIMBER, MATTHEW
APPLICANT: VALLEE, FRANCOIS
TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA
FILE REFERENCE: IPT-261.01
CURRENT APPLICATION NUMBER: US/10/967,671

PRIOR FILING DATE: 2004-10-18

PRIOR APPLICATION NUMBER: PCT/CA03/00714

PRIOR FILING DATE: 2003-05-21

PRIOR APPLICATION NUMBER: 60/382,443

; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 203
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Oryzctolagus cuniculus
US-10-055-877-203

Query Match 34.8%; Score 39; DB 6; Length 687;
Best Local Similarity 53.3%; Pred. No. 66;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 8 SLAGICGSAHLVQQL 22
Db 589 TLVGLVERHLVQAL 603

RESULT 9

US-11-079-139-12
; Sequence 12, Application US/11079139
; Publication No. US20060009620A1
; GENERAL INFORMATION:
; APPLICANT: Woolfson, Derek
; APPLICANT: Washaw, John
; APPLICANT: Pandya, Maya
; APPLICANT: Colyer, John
; TITLE OF INVENTION: PROTEIN STRUCTURES AND PROTEIN FIBRES
; FILE REFERENCE: 000487.00012
; CURRENT APPLICATION NUMBER: US/11/079,139
; PRIOR FILING DATE: 2005-03-15
; PRIOR APPLICATION NUMBER: US/10/088,417
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: PCT/GB00/03576
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB9922013.9
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptides
US-11-079-139-12

Query Match 33.9%; Score 38; DB 7; Length 29;
Best Local Similarity 54.5%; Pred. No. 3.1;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 13 COSAHLVQQLA 23
Db 8 CKNAHLKOEIA 18

RESULT 10

US-11-019-711-87
; Sequence 87, Application US/11019711
; Publication No. US20060009634A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsebrook II, John P
; APPLICANT: Tcherenev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Paturajan, Meera
; APPLICANT: Grose, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Edinger, Shlomit R

; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: Malpankar, Uriel M
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Anderson, David W
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Taudier Jr, Raymond J
; APPLICANT: Miller, Charles E
; APPLICANT: Eisen, Andrew J
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-235
; CURRENT APPLICATION NUMBER: US/11/019,711
; PRIOR FILING DATE: 2004-12-21
; PRIOR APPLICATION NUMBER: US/10/037,417
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 87
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-019-711-87

Query Match 33.9%; Score 38; DB 7; Length 307;
Best Local Similarity 42.9%; Pred. No. 40;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 HLITSLAGICQSA 16
Db 215 HLVSALIGFCATA 228

RESULT 11

US-11-019-711-88
; Sequence 88, Application US/11019711
; Publication No. US20060009634A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsebrook II, John P
; APPLICANT: Tcherenev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Paturajan, Meera
; APPLICANT: Grose, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Sciore, Paul

APPLICANT: Ellerman, Karen
APPLICANT: Malyanek, Uriel M
APPLICANT: Rothenberg, Mark
APPLICANT: Stone, David J
APPLICANT: Boldog, Ferenc L
APPLICANT: Guo, Xiaojia
APPLICANT: Shenoy, Suresh G
APPLICANT: Anderson, David W
APPLICANT: Padigaru, Muralidhara
APPLICANT: Teupler Jr, Raymond J
APPLICANT: Miller, Charles E
APPLICANT: Eisen, Andrew J
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-235
CURRENT APPLICATION NUMBER: US/11/019,711
CURRENT FILING DATE: 2004-12-21
PRIOR APPLICATION NUMBER: US/10/037,417
PRIOR FILING DATE: 2002-09-20
PRIOR APPLICATION NUMBER: 60/260,018
PRIOR FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: 60/260,360
PRIOR FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: 60/272,411
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/272,817
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/291,186
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: 60/303,231
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/305,060
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: 60/318,405
PRIOR FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: 60/318,700
PRIOR FILING DATE: 2001-09-12
NUMBER OF SEQ ID NOS: 227
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 88
LENGTH: 307
TYPE: PRT
ORGANISM: Homo sapiens
US-11-019-711-88

Query Match 33.9%; Score 38; DB 7; Length 307;
Best Local Similarity 42.9%; Pred. No. 40;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 HLITSLAGICQSA 16
DB 215 HLVSALIAIGFCATA 228

RESULT 12
US-11-179-977-15
Sequence 15, Application US/11/179,977
Publication No. US20050249789A1
GENERAL INFORMATION:
APPLICANT: Genencor International, Inc.
TITLE OF INVENTION: Alpha/Beta Hydrolyase-Fold Enzymes
FILE REFERENCE: GC511-PCT
CURRENT APPLICATION NUMBER: US/11/179,977
CURRENT FILING DATE: 2005-07-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 256
TYPE: PRT
ORGANISM: Bacillus
US-11-179-977-15

Query Match 33.5%; Score 37.5; DB 7; Length 256;
Best Local Similarity 29.2%; Pred. No. 40;

Matches 7; Conservative 9; Mismatches 5; Indels 3; Gaps 1;
QY 2 HLITSL--AGICQSAHLVQOL 22
DB 62 VHVIAISAGCPGICFASHYPERV 85

RESULT 13
US-11-033-039-1145
Sequence 1145, Application US/11/033,039
Publication No. US20060002947A1
GENERAL INFORMATION:
APPLICANT: HUMPHREYS, ROBERT
APPLICANT: XU, MINZHEN
TITLE OF INVENTION: LI-KEY/ANTIGENIC EPIOTOPE HYBRID PEPTIDE VACCINES
FILE REFERENCE: REH-2017US01
CURRENT APPLICATION NUMBER: US/11/033,039
CURRENT FILING DATE: 2005-01-11
PRIOR APPLICATION NUMBER: 10/245,871
PRIOR FILING DATE: 2002-09-17
PRIOR APPLICATION NUMBER: 10/197,000
PRIOR FILING DATE: 2002-07-17
PRIOR APPLICATION NUMBER: 09/396,813
PRIOR FILING DATE: 1999-09-14
NUMBER OF SEQ ID NOS: 1452
SOFTWARE: Patentin version 3.3
SEQ ID NO 1145
LENGTH: 32
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-11-033-039-1145

Query Match 33.0%; Score 37; DB 7; Length 32;
Best Local Similarity 61.5%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 10 AGICQSAHLVQOL 22
DB 2 AGICASVHTVSL 14

RESULT 14
US-10-793-626-370
Sequence 370, Application US/10/793,626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 370
LENGTH: 143
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-370

Query Match 33.0%; Score 37; DB 6; Length 143;
Best Local Similarity 36.8%; Pred. No. 25;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 4 LITSLAGICQSAHLVQOL 22

Db 22 ITTILMDALQNTHTQLL 40

RESULT 15

US-10-793-626-906
; Sequence 906, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 906
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-906

Query Match 33.0%; Score 37; DB 6; Length 462;
Best Local Similarity 72.7%; Pred. No. 91;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 IHLITSLAGI 12
||| |||
Db 277 IHLITSLVGI 287

Search completed: January 23, 2006, 11:53:35
Job time : 5.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-5
Sequence: 1 KIHLLTSLAGICQSAHLVQOLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	112	100.0	211 2 D84960	hypothetical prote
2	80	71.4	208 2 AH0199	conserved hypothet
3	80	71.4	213 2 S19211	ycfc protein - Esc
4	80	71.4	213 2 D90829	hypothetical prote
5	80	71.4	213 2 B85687	hypothetical prote
6	77	68.8	215 2 AB0647	conserved hypothet
7	72	64.3	205 2 I64155	hypothetical prote
8	58	51.8	205 2 B82337	conserved hypothet
9	47	42.0	847 2 G95843	conserved hypothet
10	45	40.2	302 2 S52584	crfb protein - Ew
11	45	40.2	345 2 A12217	low specificity L-
12	43	38.4	161 2 S09661	hypothetical prote
13	43	38.4	272 2 T36181	probable hydrolase
14	43	38.4	309 2 AP1171	conserved hypothet
15	43	38.4	431 2 G70561	hypothetical prote
16	43	38.4	459 2 S59612	probable membrane
17	43	38.4	3587 2 I40486	surfactin syntheta
18	42	37.5	255 2 AH2241	hypothetical prote
19	42	37.5	297 2 T26464	hypothetical prote
20	42	37.5	309 2 AH1528	conserved hypothet
21	42	37.5	346 2 A82971	low specificity 1-
22	42	37.5	359 2 D97651	hypothetical prote
23	42	37.5	336 2 S34963	ribx protein - Shi
24	42	37.5	567 2 JCS538	Rab geranylgeranyl
25	42	37.5	790 2 AE2203	adenyl-transferrase
26	42	37.5	947 2 AH0891	PI08.8 protein -
27	42	37.5	1139 2 S44800	vitellogenin precu
28	42	37.5	1677 2 T43021	uteroglobin precu
29	41.5	37.1	91 1 UGRB	

30	41.5	37.1	91	1	UGRB	uteroglobin precu
31	41	36.6	113	2	P91196	hypothetical prote
32	41	36.6	113	2	B86043	unknown protein en
33	41	36.6	116	2	H69338	hypothetical prote
34	41	36.6	139	2	H89842	hypothetical prote
35	41	36.6	192	2	I64081	glpG protein homol
36	41	36.6	294	2	AE1158	oxidoreductase hom
37	41	36.6	294	2	AB1517	oxidoreductase hom
38	41	36.6	294	2	A35266	hypothetical prote
39	41	36.6	365	2	F71437	probable resistanc
40	41	36.6	420	2	B86065	hypothetical prote
41	41	36.6	420	2	F65182	UDP-Mannac dehydro
42	41	36.6	420	2	H91218	hypothetical prote
43	41	36.6	509	2	S45413	probable membrane
44	41	36.6	560	2	AB0752	flagellar basal-bo
45	41	36.6	560	2	D30930	flagellar basal bo

ALIGNMENTS

RESULT 1

D84960 hypothetical protein [imported] - Buchnera sp. (strain APS)

C/Species: Buchnera sp. #sequence revision 02-Mar-2001 #text_change 02-Mar-2001

C/Date: 02-Mar-2001

C/Accession: D84960

R/Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakai, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A/Rfile: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. j

A/Reference number: AB4930; PMID:20445173; PMID:10993077

A/Accession: D84960

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-211 <STO>

A/Cross-references: UNIPARC:UP10000584FF; GB:AP000398; GSPDB:GN00144

A/Experimental source: strain APS

C/Genetics:

A/Gene: ycfC; BU262

Query Match 100.0%; Score 112; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIHLLTSLAGICQSAHLVQOLA 23
DB 3 KIHLLTSLAGICQSAHLVQOLA 25

RESULT 2

conserved hypothetical protein YP01637 [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C/Accession: AH0199

A/ParKhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,

Nature 413, 523-527, 2001

A/Rfile: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; PMID:21470413; PMID:11586360

A/Accession: AH0199

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-208 <KUR>

A/Cross-references: UNIPROT:Q82PQ6; UNIPARC:UP10000CD828; GB:AL590842; PIDN:CA90459.1

C/Genetics:

Query Match 71.4%; Score 80; DB 2; Length 208;
Best Local Similarity 89.5%; Pred. No. 2.6e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 ITLSIAGICQSAHLVQOLA 23
|||:|||||
Db 8 ITLALAGICQSAHLVQOLA 26

RESULT 3

ycfc protein - Escherichia coli (strain K-12)

C/Species: Escherichia coli
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: S19211; A64858
R/Green, S.M.; Drabble, W.T.
submitted to the EMBL Data Library, May 1991
A/Description: Molecular analysis of the purB-phop region of Escherichia coli K12.
A/Reference number: S19210
A/Accession: S19211
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-213 <GRB>
A/Cross-references: UNIPROT:P25746; UNIPARC:UPI000013A673; EMBL:X59307; NID:g142582; PIDN
A/Experimental source: strain K-12
R/Baltner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.U.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: A64858
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-213 <BLAT>
A/Cross-references: UNIPARC:UPI000013A673; GB:AE000213; GB:U00096; NID:g1787371; PIDN:AA
A/Experimental source: strain K-12, substrain MG1655
C/Genetics:
A/Gene: yfcC
A/Start codon: GTG

Query Match 71.4%; Score 80; DB 2; Length 213;
Best Local Similarity 89.5%; Pred. No. 2.6e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 ITLSIAGICQSAHLVQOLA 23
|||:|||||
Db 8 ITLALAGICQSAHLVQOLA 26

RESULT 4

hypothetical protein EC81604 [imported] - Escherichia coli (strain O157:H7, substrain R1

C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C/Accession: D90829
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A/Reference number: A99629; MUID:21156231; PMID:11258796
A/Accession: D90829
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-213 <HAV>
A/Cross-references: UNIPROT:Q8X736; UNIPARC:UPI00000D0AD2; GB:BA000007; PIDN:BA835027.1;
C/Genetics:
A/Experimental source: strain O157:H7, substrain R1MD 0509952
A/Gene: EC81604

Query Match 71.4%; Score 80; DB 2; Length 213;
Best Local Similarity 89.5%; Pred. No. 2.6e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 ITLSIAGICQSAHLVQOLA 23
|||:|||||
Db 8 ITLALAGICQSAHLVQOLA 26

RESULT 5

hypothetical protein yfcC [imported] - Escherichia coli (strain O157:H7, substrain EDL93:

C/Species: Escherichia coli
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C/Accession: B85687
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.U.; Mayhew
Hiler, L.; Grobeck, B.U.; Davis, N.W.; Lim, A.; Dimalante, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: B85687
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-213 <STO>
A/Cross-references: UNIPROT:Q8X736; UNIPARC:UPI00000D0AD2; GB:AE005174; NID:g12514786; P
A/Experimental source: strain O157:H7, substrain EDL933
C/Genetics:
A/Gene: yfcC

Query Match 71.4%; Score 80; DB 2; Length 213;
Best Local Similarity 89.5%; Pred. No. 2.6e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 ITLSIAGICQSAHLVQOLA 23
|||:|||||
Db 8 ITLALAGICQSAHLVQOLA 26

RESULT 6

conserved hypothetical protein STY1273 [imported] - Salmonella enterica subsp. enterica :

C/Species: Salmonella enterica subsp. enterica serovar Typhi
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AB0647
R/Parikh, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moute, S.; O'Garra, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AB0647
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-215 <PAR>
A/Cross-references: UNIPARC:UPI000059F77; GB:AL513382; PIDN:CAD08357.1; PID:g16502402;
C/Genetics:
A/Gene: STY1273

Query Match 68.8%; Score 77; DB 2; Length 215;
Best Local Similarity 84.2%; Pred. No. 8.1e-05;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 5 ITLSIAGICQSAHLVQOLA 23
|||:|||||
Db 10 ITLALAGICQSAHLVQOLA 28

RESULT 7

hypothetical protein HI0638 - Haemophilus influenzae (strain Rd KW20)

C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C/Accession: I64155
R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; FINE, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A/Reference number: A64000; MUID:95350630; PMID:7542800
 A/Accession: I64155
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-205 <TIGR>
 A/Cross-references: UNIPROT:P4796; UNIPARC:UPI000013A674; GB:U32747; GB:J42023; NID:915
 A/Note: best homolog was a hypothetical protein from *Bacterioides coli*

Query Match 64.3%; Score 72; DB 2; Length 205;
 Best Local Similarity 71.4%; Pred. No. 0.00049;
 Matches 15; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

3 HLITSLAGICQSAHLVQOLA 23
 DB 5 HDIVLALGVCSAKLVHQLA 25

RESULT 8

conserved hypothetical protein VC1127 [Imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
 C/Species: *Vibrio cholerae*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: B82237
 C/Residues: 1-205 <TIGR>
 C/Status: preliminary
 A/Residues: 1-205 <TIGR>
 A/Cross-references: UNIPROT:Q9KX9; UNIPARC:UPI00000C2B84; GB:AE004193; GB:AE003852; NID:1
 A/Note: The complete sequence of the 1.683-kb *psymb* megaplasmid from the N2-fixing endo
 A/Experimental source: serogroup O1, strain N16961; biotype El Tor
 C/Genetics:
 A/Map position: 1

Query Match 51.8%; Score 58; DB 2; Length 205;
 Best Local Similarity 61.1%; Pred. No. 0.008;
 Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

6 TLISLAGICQSAHLVQOLA 23
 DB 9 TIAFAGICQSAHLVQOLA 26

RESULT 9

conserved hypothetical protein SMD20012 [Imported] - *Sinorhizobium meliloti* (strain 1021)
 C/Species: *Sinorhizobium meliloti*
 C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C/Accession: G95843
 C/Residues: 1-847 <KUR>
 C/Status: preliminary
 A/Residues: 1-847 <KUR>
 A/Cross-references: UNIPROT:Q92X87; UNIPARC:UPI00000C83A3; GB:AL591985; PIDN:CAC48415.1;
 A/Experimental source: strain 1021, megaplasmid *psymb*
 R/Gallier, P.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federle, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kise, E.; Komp, C.; Lelaure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A/Reference number: A96039; MUID:21368234; PMID:11474104
 A/Contents: annotation

C/Genetics:
 A/Genome: SMD20012
 A/Genome: plasmid

Query Match 42.0%; Score 47; DB 2; Length 847;
 Best Local Similarity 42.3%; Pred. No. 19;
 Matches 11; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

2 HLITSLAGI---COSAHVQOLA 23
 DB 168 IRTATRLAGVDEDCRAFLAQOLA 193

RESULT 10

crb protein - *Erwinia herbicola*
 C/Species: *Erwinia herbicola*
 C/Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
 C/Accession: S52584
 C/Residues: 1-302 <LIN>
 C/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Residues: 1-302 <LIN>
 A/Cross-references: UNIPROT:Q47842; UNIPARC:UPI00000B2CFE; EMBL:M50638; NID:9148393; PTL
 A/Note: The nucleotide sequence was submitted to the EMBL Data Library, April 1992
 C/Suprafamily: geranyltransferase

Query Match 40.2%; Score 45; DB 2; Length 302;
 Best Local Similarity 40.0%; Pred. No. 15;
 Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

2 HLITSLAGICQSAHLVQOLA 21
 DB 269 IRLASHLASACONGHSTQQ 288

RESULT 11

low specificity L-threonine aldolase [Imported] - *Neotoc sp.* (strain PCC 7120)
 C/Species: *Neotoc sp.* PCC 7120
 A/Note: *Neotoc sp.* strain PCC 7120 is a synonym of *Anabaena sp.* strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C/Accession: A12217
 C/Residues: 1-345 <KUR>
 C/Status: preliminary
 A/Residues: 1-345 <KUR>
 A/Cross-references: UNIPROT:Q8YR26; UNIPARC:UPI00000C8737; GB:BA000019; PIDN:BA874995.1.
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Genome: A12217

Query Match 40.2%; Score 45; DB 2; Length 345;
 Best Local Similarity 64.3%; Pred. No. 18;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

6 TLISLAGICQSAHLV 19
 DB 69 SUSLALDCSYMSV 82

RESULT 12

hypothetical protein A2 (insertion sequence ISR1) - *Rhizobium sp.*

C/Species: *Rhizobium* sp.
C/Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004
C/Accession: S09661
R/Priifer, U.B.; Kallinowski, J.; Rueger, B.; Heumann, W.; Puhler, A.
A/Title: IS1, a transposable DNA sequence resident in *Rhizobium* class IV strains, shows
A/Reference number: S09659; MUID:89297222; PMID:2544911
A/Accession: S09661
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-161 <PRI>
A/Cross-references: UNIPROT:P17984; UNIPARC:UPI000013B359; EMBL:X06616; MID:948670; PIDN
C/Genetics:
A/Genome: insertion sequence
A/Start codon: GTG

Query Match 38.4%; Score 43; DB 2; Length 161;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 10 AGICQSAHLVQOLA 23
Db 118 AGLCRSARTRQPA 131

RESULT 13
T36181
probable hydrolase - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T36181
R/Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, March 1999
A/Reference number: Z21600
A/Accession: T36181
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-272 <SAU>
A/Cross-references: UNIPROT:Q92428; UNIPARC:UPI00000DAF2A; EMBL:AL035707; PIDN:CAM38877.
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCODB:SCB29.0
C/Superfamily: carboxyl esterase

Query Match 38.4%; Score 43; DB 2; Length 272;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 9 LAGICQSAHLVQOLA 23
Db 97 VAGISMGAIHQBELA 111

RESULT 14
AF1171
conserved hypothetical protein Imo0774 [imported] - *Listeria monocytogenes* (strain EGD-e
C/Species: *Listeria monocytogenes*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AF1171
R/Glasser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Kars, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me
ok, C.; Schlueter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A/Title: Comparative genomes of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AF1171
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-309 <GLA>
A/Cross-references: UNIPROT:Q8YWM8; UNIPARC:UPI0000054D35; GB:NC_003210; PIDN:CA098852.1
A/Experimental source: strain EGD-e

C/Genetics:
A/Gene: Imo0774

Query Match 38.4%; Score 43; DB 2; Length 309;
Best Local Similarity 33.3%; Pred. No. 34;
Matches 7; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IHITLSLAGICQSAHLVQOL 22
Db 218 LHITLTKGLDPAAMWIPOL 238

RESULT 15
G70561
hypothetical protein Rv3630 - *Mycobacterium tuberculosis* (strain H37Rv)
C/Species: *Mycobacterium tuberculosis*
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004
C/Accession: G70561
R/Cole, S.T.; Broese, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Suleron, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: G70561
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-431 <COL>
A/Cross-references: UNIPROT:O06377; UNIPARC:UPI0000139264; GB:Z95436; GB:AL123456; MID:G
A/Experimental source: strain H37Rv
C/Genetics:
A/Gene: Rv3630
C/Superfamily: uncharacterized conserved protein

Query Match 38.4%; Score 43; DB 2; Length 431;
Best Local Similarity 40.9%; Pred. No. 46;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 2 IHITLSLAGICQSAHLVQOLA 23
Db 143 VALISGLAGFCPLHATLGMILA 164

Search completed: January 23, 2006, 09:32:50
Job time : 10.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 Seconds
(without alignments)
267.334 Million cell updates/sec

Title: US-10-501-838A-5
Perfect score: 112
Sequence: 1 KIHLLITSLAGICQSAHLVQQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Uniprot_05.80+
1: uniprot_sprot+
2: uniprot_trembl+*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	211	1 Y262_BUCAI	P57350 buchera ap
2	88	78.6	212	2 Q6D4E8_ERWCT	Q6D4E8 erwina car
3	80	71.4	208	1 Y1637_YERPE	Q8ZFE6 yerinia pe
4	80	71.4	208	2 Q6E8Q3_YERPS	Q6E8Q3 yerinia ps
5	80	71.4	213	1 YCPC_ECO57	Q8X736 escherichia
6	80	71.4	213	1 YCPC_ECOL6	Q8E1B7 escherichia
7	80	71.4	213	1 YCPC_ECOL1	P25746 escherichia
8	80	71.4	213	1 YCPC_SHIFL	Q831F8 shigella fl
9	77	68.8	213	1 YCPC_SALTY	Q8Z7D0 salmonella
10	77	68.8	213	1 YCPC_SALTY	Q8Z7D0 salmonella
11	77	68.8	213	1 YCPC_SALTY	Q8Z7D0 salmonella
12	73	65.1	204	2 Q6SVV5_MANSW	Q6SVV5 manheimia
13	73	65.1	208	1 Y2805_PHOHL	Q7N3D4 photorhabdu
14	73	65.2	215	2 Q57QCI_SALCH	O57QCI salmonella
15	72	64.3	205	1 Y638_HABIN	P44796 haemophilus
16	72	64.3	205	2 Q4QMS9_HABE8	Q4QMS9 haemophilus
17	71	63.4	203	1 Y1850_PASNU	Q9CJY8 pasteurella
18	65	58.0	215	1 Y1650_HABDU	Q7V135 haemophilus
19	61	54.5	217	1 Y243_BUCBP	Q89AM4 buchera ap
20	60	53.6	205	2 Q5E3W6_VIBP1	Q5E3W6 vibrio fusc
21	58	51.8	205	1 Y1127_VIBCH	Q9KXK9 vibrio chol
22	58	51.8	205	1 Y1129_VIBPA	Q87QM0 vibrio para
23	56	50.0	205	1 Y1342_VIBBY	Q7M1C5 vibrio vuln
24	56	50.0	205	1 Y2927_VIBVU	Q8DBP4 vibrio vuln
25	53	47.3	207	2 Q60CAB_METCA	Q60CAB methylococc
26	49	43.8	426	2 Q9XDF7_STRGR	Q9XDF7 streptomyce
27	49	43.8	987	2 Q53MD2_ORYSA	Q53MD2 oryza sativ
28	47	42.0	34	2 Q7VLM0_HABDU	Q7VLM0 haemophilus
29	47	42.0	199	2 Q8GQ40_TSEAB	Q8GQ40 pseudomonas
30	47	42.0	847	2 Q92XK7_RHINS	Q92XK7 rhizobium m
31	46	41.1	205	2 Q6L719_PHOBR	Q6L719 photobacter

ALIGNMENTS

32	46	41.1	355	2	Q51Z16_MAGGR	Q51Z16 magnaporthe
33	46	41.1	1025	2	Q6AUZ9_ORYSA	Q6AUZ9 oryza sativ
34	46	41.1	1083	2	Q7NCL2_GROVA	Q7NCL2 gloeobacter
35	46	41.1	2565	2	Q4IHG8_GIBZB	Q4IHG8 gibberella
36	45	40.2	184	2	Q44D8_MACFA	Q44D8 macaca faec
37	45	40.2	224	2	Q41WN3_AZOVI	Q41WN3 azotobacter
38	45	40.2	286	2	Q5R9F1_PONPY	Q5R9F1 pongo pygma
39	45	40.2	302	2	Q47842_ENTAG	Q47842 enterobacte
40	45	40.2	330	1	MERG_HUMAN	Q9NKK6 homo sapien
41	45	40.2	345	2	Q8YRZ6_ANASP	Q8YRZ6 anabaena sp
42	45	40.2	447	2	Q4K6R5_PSEFS	Q4K6R5 pseudomonas
43	45	40.2	486	2	Q74DW2_GEOSL	Q74DW2 geobacter s
44	45	40.2	508	2	Q5BAG2_EMENT	Q5BAG2 aspergillus
45	45	40.2	1102	2	Q4SDJ8_TETNG	Q4SDJ8 tetradon n

RESULT 1

AC	P57350;	STANDARD;	PRT;	211 AA.
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Hypothetical UPF0274 protein B0262.			
GN	OrderedLocustNames=B0262;			
OS	Buchera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum			
OS	synthetic bacterium).			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Buchera.			
OX	NCBI_Taxid=118099;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].			
RC	STRAIN=Tokyo 1998;			
RX	MEDLINE=20445173; PubMed=10993077; DOI=10.1038/35024074;			
RA	Shigenobu S., Matsumbe H., Hattori M., Sakaki Y., Ishikawa H.;			
RT	"Genome-sequence of the endocellular bacterial symbiont of aphids			
RT	Buchera sp. APS."			
RL	Nature 407:81-86(2000).			
CC	-1- SIMILARITY: Belongs to the UPF0274 family.			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use as long as its content is in no way modified and this statement is not			
CC	removed.			
CC	EMBL; BA000003; BAB12972.1; -; Genomic_DNA.			
DR	HAMAP; MF_00695; -; 1.			
DR	InterPro; IPR007451; DUF489.			
DR	Pfam; PF04356; DUF489; 1.			
KW	Complete proteome; Hypothetical protein.			
SQ	SEQUENCE 211 AA; 24303 MW; CA9C95140E478A5C CRC64;			
QY	Query Match	100.0%;	Score 112;	DB 1;
QY	Best local Similarity	100.0%;	Pred. No. 8.1e-10;	Length 211;
QY	Matches	23;	Conservative	0;
QY			Mismatches	0;
QY			Indels	0;
QY			Gaps	0;
DB	1 KIHLLITSLAGICQSAHLVQQLA 23			
DB	3 KIHLLITSLAGICQSAHLVQQLA 25			
RESULT 2				
AC	Q6D4E8_ERWCT			
ID	Q6D4E8; ERWCT PRELIMINARY;			
IC	Q6D4E8; ERWCT PRELIMINARY;			
DT	25-OCT-2004 (TrEMBLrel. 28, Created)			
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)			
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)			
DE	Hypothetical protein.			

```
GN OrderedLocusNames=ECA2443;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
CX NCBI_TaxID=29471;
RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebailia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holleva M.C., Thomson N.R., Bentley S.D., Churcher L.V.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Nordczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmon D.G.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
DR EMBL; BX950851; CAG75345.1; -; Genomic_DNA.
DR SMR; Q6D488; 2-207.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 212 AA; 22892 MW; A67813005F32B150 CRC64;

Query Match 78.6%; Score 88; DB 2; Length 212;
Best Local Similarity 94.7%; Pred. No. 6.8e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 ITL5IAGICQSAHLVQQLA 23
Db 8 ITL5IAGICQSAHLVQQLA 26

RESULT 3
Y1637 YERPPE STANDARD; PRT; 208 AA.
AC Q8ZFQ6; Q8D001;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein YP01637/Y1798/YP1767.
GN OrderedLocusNames=YP01637, Y1798, YP1767;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
CX NCBI_TaxID=632;
RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebailia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Baskham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RT Nature 413:523-527(2001).
RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
DOI=10.1126/SCIENCE.116.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Heuvelink J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McConough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Berry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RT J. Bacteriol. 184:4601-4611(2002).
```

```
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans."
RL DNA Res. 11:179-197(2004).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AJ141419; CAC90459.1; -; Genomic DNA.
DR EMBL; AE013782; AAM85366.1; ALT_INIT; Genomic DNA.
DR EMBL; AE017133; AAS61994.1; -; Genomic DNA.
DR PIR; AH0199; AH0199.
DR SMR; Q8ZFQ6; 2-207.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
KT CONFLICT 1 M -> MIV (in Ref. 2).
SQ SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE499723 CRC64;

Query Match 71.4%; Score 80; DB 1; Length 208;
Best Local Similarity 89.5%; Pred. No. 0.00014;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITL5IAGICQSAHLVQQLA 23
Db 8 ITL5IAGICQSAHLVQQLA 26

RESULT 4
Q66903 YERPS PRELIMINARY; PRT; 208 AA.
AC Q66903;
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=YPTB2431;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
CX NCBI_TaxID=633;
RN RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP2953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chait N.S.G., Carniel E., Latimer F.W., Lanerain J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Decheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia B.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH21669.1; -; Genomic DNA.
DR SMR; Q66903; 2-207.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE499723 CRC64;

Query Match 71.4%; Score 80; DB 2; Length 208;
```

Best Local Similarity 89.5%; Pred. No. 0.00014;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITSLAGICOSARLVQOLA 23
DB 8 ITLALAGICOSARLVQOLA 26

RESULT 5

YCF_C_ECO57

ID YCF_C_ECO57 STANDARD; PRT; 213 AA.

AC Q8X736;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocustNames=21861, Ecgi604;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=0157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grobeck C.J., Davis N.W., Lim A., Dialante E.T., Potomousis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RT Nature 409:529-533 (2001).

RL Nature 409:529-533 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=0157:H7 / Sakai / RMD 050952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hatori M., Shingawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22 (2001).

CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL, AB005174; AGS55958.1; -, Genomic DNA.
CC EMBL, BA000007; BAB55027.1; -, Genomic DNA.
CC PIR, B85687; B85687.
CC PIR, D90829; D90829.
CC SWR, Q8X736; 2-213.
CC HAMAP, MF_00695; -1.
CC InterPro, IPR007451; DUF489.
CC Pfam, PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22947 MW; E2EP9678C3844E2 CRC64;

Query Match 71.4%; Score 80; DB 1; Length 213;
Best Local Similarity 89.5%; Pred. No. 0.00014;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITSLAGICOSARLVQOLA 23
DB 8 ITLALAGICOSARLVQOLA 26

RESULT 6

YCF_C_ECOL6
ID YCF_C_ECOL6 STANDARD; PRT; 213 AA.

AC Q8F1B7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocustNames=c1511;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=06:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
Rasko D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).

CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL, AE016759; AA079980.1; ALT_INIT; Genomic DNA.
CC SMR, Q8F1B7; 2-213.
CC HAMAP, MF_00695; -1.
CC InterPro, IPR007451; DUF489.
CC Pfam, PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22934 MW; 71C9F96F88AA37BB CRC64;

Query Match 71.4%; Score 80; DB 1; Length 213;
Best Local Similarity 89.5%; Pred. No. 0.00014;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 ITSLAGICOSARLVQOLA 23
DB 8 ITLALAGICOSARLVQOLA 26

YCF_C_ECOL1
ID YCF_C_ECOL1 STANDARD; PRT; 213 AA.

AC P25746;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC (ORF-23).
GN Name=ycfC; OrderedLocustNames=b1132;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12;
RA Green S.M., Drabble W.T.;
RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.

RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12;
RX MEDLINE=92104952; PubMed=1729205;
RA He B., Smith J.M., Zaikin H.;
RT "Escherichia coli purB gene: cloning, nucleotide sequence, and
regulation by purR.";

```

RL J. Bacteriol. 174:130-136(1992).
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Verde J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirtpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905322;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiiuchi T.;
RT "A 718-Kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to
frameshifts.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; X59307; CAA41995.1; -; Genomic_DNA.
DR EMBL; M74924; -; NOT ANNOTATED CDS; Genomic_DNA.
DR EMBL; U00096; AAC74216.1; -; Genomic_DNA.
DR EMBL; D90748; BA35954.1; -; Genomic_DNA.
DR EMBL; D90749; BA35963.1; -; Genomic_DNA.
DR PIR; S19211; S19211.
DR PDB; 1QZ4; X-ray; A=2-213.
DR PDB; 1SDI; X-ray; A=2-213.
DR ECHOBASE; EB1321; -.
DR EcoGene; EG1345; yjcF.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
DR 3D-structure; Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22948 MW; E2EEF9698C3CBE42 CRC64;

Query Match 71.4%; Score 80; DB 1; Length 213;
Best Local Similarity 89.5%; Pred. No. 0.00014;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 ITSLAGICQSAHLVQOLA 23
| | | | | | | | | | | | | | | | | | | | |
Db 8 ITLALAGICQSAHLVQOLA 26

RESULT 8
YCFC SHIFL
ID YCFC SHIFL STANDARD; PRT; 213 AA.
AC Q83LE8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein yjcF.
GN Name=yjcF; OrderedLocuNames=SF1151, S1234;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_Taxid=623;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang Y., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; AB005674; AAN42768.1; ALT INIT; Genomic_DNA.
DR EMBL; AE016982; AAP16657.1; -; Genomic_DNA.
DR SRR; Q83LF8; 2-213.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22894 MW; 0B0BEC698C29FB4D CRC64;

Query Match 71.4%; Score 80; DB 1; Length 213;
Best Local Similarity 89.5%; Pred. No. 0.00014;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 ITSLAGICQSAHLVQOLA 23
| | | | | | | | | | | | | | | | | | | | |
Db 8 ITLALAGICQSAHLVQOLA 26

RESULT 9
YCFC SALTI
ID YCFC SALTI STANDARD; PRT; 213 AA.
AC Q827H0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein yjcF.
GN Name=yjcF; OrderedLocuNames=STY1273, t1687;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_Taxid=601;
RN [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies K.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jags K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,

```

```

RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typh1 CT18."
RL Nature 413:848-852(2001).
RN (2)
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typh1 strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AL627269; CAD08357.1; ALT_INIT; Genomic DNA.
CC EMBL; AE016839; AA069312.1; ALT_INIT; Genomic DNA.
CC SMR; Q827H0; 2-213.
CC DR HAMAP; MF_00695; -; 1.
CC DR InterPro; IPR007451; DUF489.
CC DR Pfam; PF04356; DUF489; 1.
CC DR Complete proteome; Hypothetical protein.
CC KW SEQUENCE 213 AA; 22942 MW; 63192DIC8066ED58 CRC64;
SQ
Query Match 68.8%; Score 77; DB 1; Length 213;
Best Local Similarity 84.2%; Pred. No. 0.00043;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 5 ITSLAGICQSARLVQOLA 23
Db 8 ITLALSGICQSARLVQOLA 26
RESULT 10
YCFC SALTY STANDARD; PRT; 213 AA.
AC 082P25;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein yCFC.
GN Name=yCFC; OrderedLocusNames=STM1233;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN (1)
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.R., Spieth J., Clifton S.W., Latreille P.,
RA Courthey L., Porwollik S., Ali J., Dante M., Du R., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan B., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

```

```

DR EMBL; AE008754; AAL20162.1; ALT_INIT; Genomic DNA.
DR SMR; Q82P25; 2-213.
DR StyGene; SGR777?; yCFC.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
DR Complete proteome; Hypothetical protein.
DR KW SEQUENCE 213 AA; 22916 MW; 63192DIC807C5D58 CRC64;
SQ
Query Match 68.8%; Score 77; DB 1; Length 213;
Best Local Similarity 84.2%; Pred. No. 0.00043;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 5 ITSLAGICQSARLVQOLA 23
Db 8 ITLALSGICQSARLVQOLA 26
RESULT 11
QSPMU3 SALPA PRELIMINARY; PRT; 213 AA.
AC 05PMU3;
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Hypothetical protein yCFC.
GN Name=yCFC; OrderedLocusNames=SPAL617;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN (1)
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=ATCC 9150;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.R., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozerik P., McClelland M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlbeg S., Strong C., Du F., Carter J., Kremliki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaunt J., Fronick C., Magrini V., Nhan M., Warren W., Flores L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typh1, human-
RT restricted serovars of Salmonella enterica that cause typhoid."
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP000026; AAV77544.1; -; Genomic DNA.
DR SMR; 05PMU3; 2-213.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
DR Complete proteome; Hypothetical protein.
DR KW SEQUENCE 213 AA; 22916 MW; 63192DIC807C5D58 CRC64;
SQ
Query Match 68.8%; Score 77; DB 2; Length 213;
Best Local Similarity 84.2%; Pred. No. 0.00043;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 5 ITSLAGICQSARLVQOLA 23
Db 8 ITLALSGICQSARLVQOLA 26
RESULT 12
Q65VVS MANSM PRELIMINARY; PRT; 204 AA.
AC 065VVS;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MS0298;
OS Mannheimia succiniciproducens (strain MBE155E).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.

```

OX NCBI_TaxID=221988;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15378067; DOI=10.1038/nbt1010;
 RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
 RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;
 RT "The genome sequence of the capnophilic rumen bacterium Mannheimia
 succiniciproducens".
 RL Nat. Biotechnol. 22:1275-1281(2004).
 DR EMBL; AB016827; A036905.1; -; Genomic_DNA.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 204 AA; 23036 MW; 30d4a70a30c33c6b CRC64;

Query Match 66.1%; Score 74; DB 2; Length 204;
 Best Local Similarity 78.9%; Pred. No. 0.0013;
 Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 ITLSIAGICGSAHLYVOOLA 23
 |||:|||||:|||||
 DB 8 ITLALAGVCGSAKLVQOFA 26

RESULT 13
 Y2805 PHOLL STANDARD; PRT; 208 AA.
 AC Q7N3B4;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein plu2805.
 GN OrderedLocustNames=plu2805;
 OS Photorhabdus luminescens (subsp. laumontii).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Photorhabdus.
 OX NCBI_TaxID=141679;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=TT01;
 RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
 RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Gavaudan A.,
 RA Taourit S., Bocs S., Bourreau-Bude C., Chandler M., Charles J.-F.,
 RA Dasse B., Desose R., Derzelle S., Freysinet G., Gaudinot S.,
 RA Mediague C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,
 RA Zouine M., Glaser P., Boemare N., Danchin A., Kunet P.;
 RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
 luminescens".
 RL Nat. Biotechnol. 21:1307-1313(2003).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; BX571866; CAB15179.1; -; Genomic_DNA.
 DR Photobase; plu2805; -;
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 208 AA; 22831 MW; 3b484f91bf6b36ad CRC64;

Query Match 65.2%; Score 73; DB 1; Length 208;
 Best Local Similarity 78.9%; Pred. No. 0.0019;
 Matches 15; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 ITLSIAGICGSAHLYVOOLA 23
 |||:|||||:|||||
 DB 8 ITLALAGVCGSAKLVQOOLS 26

RESULT 14
 O570C1 SALCH PRELIMINARY; PRT; 215 AA.
 ID O570C1
 AC O570C1;
 DT 10-MAY-2005 (TRENBLUREL. 30, Created)
 DT 10-MAY-2005 (TRENBLUREL. 30, Last sequence update)
 DT 10-MAY-2005 (TRENBLUREL. 30, Last annotation update)
 DE Hypothetical protein ycfC.
 GN Name=ycfC; OrderedLocustNames=SC1184;
 OS Salmonella cholerae-suis (Salmonella enterica).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=591;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SC-B67;
 RX PubMed=15781495;
 RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
 RA Wang H.-S., Lee Y.-S.;
 RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
 RT highly invasive and resistant zoonotic pathogen".
 RL Nucleic Acids Res. 33:1690-1698(2005).
 DR EMBL; AB017220; AAY65090.1; -; Genomic_DNA.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 215 AA; 23157 MW; C153B20629FAB8BD CRC64;

Query Match 65.2%; Score 73; DB 2; Length 215;
 Best Local Similarity 83.3%; Pred. No. 0.0019;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 ITLSIAGICGSAHLYVOOL 22
 |||:|||||:|||||
 DB 10 ITLALSGICGSAHLYVOOL 27

RESULT 15
 Y638 HAEIN STANDARD; PRT; 205 AA.
 ID Y638 HAEIN
 AC P44796;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein HI0638.
 GN OrderedLocustNames=HI0638;
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Flaischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McEweney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Iu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uetzerack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.". Science 269:496-512(1995).
 RL Science 269:496-512(1995).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----

This page blank (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838a-5
Perfect score: 112
Sequence: 1 KIHLLTSLAGICQSAHLVQOLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA_Main:*

- 1: /cgn2_6/pcodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/pcodata/1/pubppaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/pcodata/1/pubppaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/pcodata/1/pubppaa/US10_PUBCOMB.pep:*
- 5: /cgn2_6/pcodata/1/pubppaa/US10_PUBCOMB.pep:*
- 6: /cgn2_6/pcodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112	100.0	23	4	US-10-665-184-5
2	112	100.0	23	5	US-10-942-300-5
3	112	100.0	23	5	US-10-501-838a-5
4	112	100.0	211	4	US-10-665-184-63
5	112	100.0	211	5	US-10-942-300-63
6	112	100.0	211	5	US-10-501-838a-57
7	80	71.4	22	4	US-10-665-184-27
8	80	71.4	22	5	US-10-942-300-27
9	80	71.4	22	5	US-10-501-838a-27
10	80	71.4	23	4	US-10-665-184-3
11	80	71.4	23	4	US-10-665-184-26
12	80	71.4	23	5	US-10-942-300-3
13	80	71.4	23	5	US-10-942-300-26
14	80	71.4	23	5	US-10-501-838a-3
15	80	71.4	23	5	US-10-501-838a-26
16	80	71.4	24	4	US-10-665-184-25
17	80	71.4	24	5	US-10-942-300-25
18	80	71.4	24	5	US-10-501-838a-25
19	80	71.4	25	5	US-10-501-838a-19
20	80	71.4	25	5	US-10-501-838a-23
21	80	71.4	26	5	US-10-501-838a-21
22	80	71.4	29	4	US-10-665-184-37
23	80	71.4	29	4	US-10-942-300-37
24	80	71.4	29	5	US-10-501-838a-18
25	80	71.4	30	5	US-10-501-838a-37
26	80	71.4	30	4	US-10-665-184-22
27	80	71.4	30	4	US-10-665-184-36

28	80	71.4	30	5	US-10-942-300-22	Sequence 22, Appl
29	80	71.4	30	5	US-10-942-300-36	Sequence 36, Appl
30	80	71.4	30	5	US-10-501-838a-20	Sequence 20, Appl
31	80	71.4	30	5	US-10-501-838a-22	Sequence 22, Appl
32	80	71.4	30	5	US-10-501-838a-36	Sequence 36, Appl
33	80	71.4	31	4	US-10-665-184-35	Sequence 35, Appl
34	80	71.4	31	5	US-10-942-300-35	Sequence 35, Appl
35	80	71.4	31	5	US-10-501-838a-35	Sequence 35, Appl
36	80	71.4	213	4	US-10-665-184-61	Sequence 61, Appl
37	80	71.4	213	5	US-10-942-300-61	Sequence 61, Appl
38	80	71.4	213	5	US-10-501-838a-55	Sequence 55, Appl
39	73	65.2	23	4	US-10-665-184-29	Sequence 29, Appl
40	73	65.2	23	5	US-10-942-300-29	Sequence 29, Appl
41	73	65.2	23	5	US-10-501-838a-29	Sequence 29, Appl
42	73	65.2	29	4	US-10-665-184-33	Sequence 33, Appl
43	73	65.2	29	5	US-10-942-300-33	Sequence 33, Appl
44	73	65.2	30	5	US-10-501-838a-33	Sequence 33, Appl
45	72	64.3	23	4	US-10-665-184-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-10-665-184-5
; Sequence 5, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Buchnera aphidicola
US-10-665-184-5
Query Match 100.0%; Score 112; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 KIHLLTSLAGICQSAHLVQOLA 23
1 KIHLLTSLAGICQSAHLVQOLA 23
RESULT 2
US-10-942-300-5
; Sequence 5, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

PRIOR FILING DATE: 2003-09-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5
LENGTH: 23
TYPE: PRT
ORGANISM: Buchnera aphidicola
US-10-942-300-5

Query Match 100.0%; Score 112; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIHILITSLAGICQSAHLVQOLA 23
Db 1 KIHILITSLAGICQSAHLVQOLA 23

RESULT 3

US-10-501-838A-5
Sequence 5, Application US/10501838A
Publication No. US20050215478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5
LENGTH: 23
TYPE: PRT
ORGANISM: Buchnera aphidicola
US-10-501-838A-5

Query Match 100.0%; Score 112; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIHILITSLAGICQSAHLVQOLA 23
Db 1 KIHILITSLAGICQSAHLVQOLA 23

RESULT 4

US-10-665-184-63
Sequence 63, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Cohen, Elnat
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501CIP
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 63
LENGTH: 211
TYPE: PRT
ORGANISM: Buchnera aphidicola

US-10-665-184-63

Query Match 100.0%; Score 112; DB 4; Length 211;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIHILITSLAGICQSAHLVQOLA 23
Db 3 KIHILITSLAGICQSAHLVQOLA 25

RESULT 5

US-10-942-300-63
Sequence 63, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-503
CURRENT APPLICATION NUMBER: US/10/942,300
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 63
LENGTH: 211
TYPE: PRT
ORGANISM: Buchnera aphidicola
US-10-942-300-63

Query Match 100.0%; Score 112; DB 5; Length 211;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIHILITSLAGICQSAHLVQOLA 23
Db 3 KIHILITSLAGICQSAHLVQOLA 25

RESULT 6

US-10-501-838A-57
Sequence 57, Application US/10501838A
Publication No. US20050215478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: US/10/501,838A
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 57
LENGTH: 211
TYPE: PRT
ORGANISM: Buchnera aphidicola
US-10-501-838A-57

Query Match 100.0%; Score 112; DB 5; Length 211;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY		1 KIHLLTSLAGICQSAAHLVQLA	23
Dd		3 KIHLITLSLAGICQSAAHLVQLA	25

```

RESULT 7
US-10-665-184-27
Sequence 27, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
APPLICANT: Cohen, Elnat
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
TITLE OF INVENTION: Biological Barrier
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatencIn version 3.2
SEQ ID NO 27
LENGTH: 22
TYPE: prt
ORGANISM: Escherichia coli
US-10-665-184-27

```

Query Match	71.4%	Score 80	DB 4	Length 22
Best Local Similarity	89.5%	Pred. NO.	5.4e-05	
Matches 17; Conservative		1; Mismatches	1;	Indels 0; Gaps 0

```

QY      5 ITLSLAGICQSAHLVQOLA 23
         |||:||||| |||
Db      4 ITLALAGICQSAHLVQOLA 22

```

```

RESULT 8
US-10-942-300-27
/ Sequence 27, Application US/10942300
/ Publication No. US20050136103A1
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Elina
/ APPLICANT: Ben-Sasson, Shmuel
/ TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
/ TITLE OF INVENTION: Biological Barrier
/ FILE REFERENCE: 24348-503
/ CURRENT APPLICATION NUMBER: US/10/942.300
/ CURRENT FILING DATE: 2004-09-16
/ PRIOR APPLICATION NUMBER: 10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 10/664,989
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 60/503,615
/ PRIOR FILING DATE: 2003-09-17
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 27
/ LENGTH: 22
/ TYPE: prt
/ ORGANISM: Escherichia coli
US-10-942-300-27

```

Query Match	71.4%	Score 80;	DB 5;	Length 22;
Best Local Similarity	89.5%	Pred. No. 5.4e-05;		
Matches 17; Conservative	1;	Mismatches 1;	Indels 0;	Gaps 0.

```

OY      5 ITLSLAGICQSAHLVQQLA 23
        |||:||||||| |||||
Db      4 ITLALAGICQSARLVQQLA 22

```

RESULT 9
US-10-501-838A-27
; Sequence 27, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; Addendum: Patent Office
; Chemical

Query Match	71.4%	Score 80:	DB 5:	length 22;
Best Local Similarity	89.5%	Pred. No.	5.4e-05;	
Matches 17; Conservative	1;	Mismatches	1;	Indels 0; Gaps 0;

```
QY      5 ITLSLAGICQSAHLVQOLA 23
      |||:|||||||  |||||
Db      4 ITLALAGICQSAHLVQOLA 22
```

```

RESULT 10
US-10-665-184-3
; Sequence 3, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Eilat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-665-184-3

```

Query Match	71.4%	Score 80;	DB 4;	length 23;
Best Local Similarity	89.5%	Pred. No.	5.6e-05;	
Matches 1/; Conservative	1;	Mismatches	1;	Indels 0; Gaps 0;

QY	5	ITLSIAGICQSAHLVQOLA	23
	:		
Db	5	ITLALAGICQSAHLVQOLA	23

RESULT 11
US-10-665-184-26
; Sequence 26, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel

```

Query Match 71.4%; Score 80; DB 5; Length 23;
Best Local Similarity 89.5%; Pred. No. 5,6e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      5 TTSLAGICGSAHLVQOLA 23
      |||:|||||:|||||
Db      5 ITTLAGICGSAHLVQOLA 23

RESULT 13
US-10-942-300-26
; Sequence 26, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saeon, Shmuel
; APPLICANT: Cohen, Eilat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-503
; CURRENT PRIORITY NUMBER: US/10/942,300

```

```

Query Match          71.4%; Score 80; DB 5; Length 23;
Best Local Similarity 89.5%; Pred. No. 5.6e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy      5      ITTSLAGICOSARLVQOLA 23
      |||:||||| |||
Db      5      ITTSLAGICOSARLVQOLA 23

RESULT 15
US-10-501-838A-26
; Sequence 26, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OR INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07

```

/ NUMBER OF SEQ ID NOS: 72
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 26
 / LENGTH: 23
 / TYPE: PRT
 / ORGANISM: Escherichia coli
 US-10-501-838a-26

Query Match 71.4%; Score 80; DB 5; Length 23;
 Best Local Similarity 89.5%; Pred. No. 5.6e-05;
 Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ITLTLAGICQSARLVQQLA 23
 |||:|||||||
 Db 5 ITLTLAGICQSARLVQQLA 23

Search completed: January 23, 2006, 11:52:32
 Job time : 60 secs

The Page Blank (uspofo)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 23, 2006, 11:32:56 ; Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838a-6

Perfect score: 108
Sequence: 1 DPROQIALGAVFESALVDKLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubppaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubppaa/US05_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubppaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	40.7	585	US-11-012-762-6	Sequence 6, Appl1
2	44	40.7	890	US-10-510-386-26	Sequence 26, Appl1
3	42	38.9	87	US-09-978-360A-472	Sequence 472, App
4	41.5	38.4	537	US-11-144-236-6	Sequence 6, Appl1
5	40	37.0	154	US-09-978-360A-766	Sequence 766, App
6	37.0	36.1	306	US-10-467-657-8116	Sequence 8116, App
7	39	36.1	753	US-10-485-517-173	Sequence 173, App
8	38	35.2	387	US-11-010-795-5	Sequence 5, Appl1
9	38	35.2	387	US-11-010-795-14	Sequence 14, Appl1
10	38	35.2	387	US-11-010-795-18	Sequence 18, Appl1
11	38	35.2	469	US-11-133-360-20	Sequence 20, Appl1
12	38	35.2	469	US-11-133-346-20	Sequence 20, Appl1
13	38	35.2	527	US-11-069-642-21	Sequence 21, Appl1
14	37.5	34.7	7968	US-11-143-980-49	Sequence 49, Appl1
15	37	34.3	125	US-10-516-768-18	Sequence 18, Appl1
16	37	34.3	161	US-10-467-657-22	Sequence 22, Appl1
17	37	34.3	161	US-10-467-657-6198	Sequence 6198, App
18	37	34.3	162	US-10-516-768-21	Sequence 21, Appl1
19	37	34.3	208	US-10-131-826A-510	Sequence 510, App
20	37	34.3	420	US-10-131-826A-260	Sequence 260, App
21	37	34.3	508	US-11-075-185-26	Sequence 26, Appl1
22	37	34.3	686	US-11-099-691-9	Sequence 9, Appl1
23	37	34.3	1073	US-10-467-657-5230	Sequence 5230, App
24	36.5	33.8	1126	US-11-110-480-7	Sequence 7, Appl1
25	36.5	33.8	1126	US-11-110-480-9	Sequence 9, Appl1

26	36.5	33.8	1126	US-11-110-480-16	Sequence 16, Appl1
27	36.5	33.8	1126	US-11-110-480-27	Sequence 27, Appl1
28	36.5	33.8	1126	US-11-110-480-37	Sequence 37, Appl1
29	36.5	33.8	1126	US-11-110-480-47	Sequence 47, Appl1
30	36.5	33.8	1126	US-11-110-480-59	Sequence 59, Appl1
31	36.5	33.8	1126	US-11-110-480-82	Sequence 82, Appl1
32	36.5	33.8	1126	US-11-110-480-88	Sequence 88, Appl1
33	36	33.3	193	US-11-186-284-157	Sequence 157, App
34	36	33.3	193	US-11-186-284-159	Sequence 159, App
35	36	33.3	238	US-10-623-155-174	Sequence 174, App
36	36	33.3	673	US-11-058-727-14	Sequence 14, Appl1
37	36	33.3	673	US-11-108-389-14	Sequence 14, Appl1
38	36	33.3	820	US-10-858-730-211	Sequence 211, App
39	36	33.3	1210	US-11-058-727-4	Sequence 4, Appl1
40	36	33.3	1210	US-11-108-389-4	Sequence 4, Appl1
41	35.5	32.9	539	US-10-514-250A-28	Sequence 28, Appl1
42	35.5	32.9	539	US-10-514-250A-29	Sequence 29, Appl1
43	35.5	32.9	539	US-10-514-250A-30	Sequence 30, Appl1
44	35.5	32.9	539	US-10-514-250A-31	Sequence 31, Appl1
45	35.5	32.9	539	US-10-514-250A-32	Sequence 32, Appl1

ALIGNMENTS

```
RESULT 1
US-11-012-762-6
; Sequence 6, Application US/11012762
; Publication No. US20050244815A1
; GENERAL INFORMATION:
; APPLICANT: Georgia State University Research Foundation, Inc.
; TITLE OF INVENTION: Compositions and Methods for Viral Resistance Gence
; FILE REFERENCE: GSU1.PCT
; CURRENT FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: PCT/US03/19300
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/390,046
; PRIOR FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-012-762-6

Query Match          40.7%; Score 44; DB 7; Length 585;
Best Local Similarity 47.4%; Pred. No. 6.7;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Cy      1 DPROQIALGAVFESALV 19
Db      163 DPROQIALGAVFESALV 181

RESULT 2
US-10-510-386-26
; Sequence 26, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26
```

```

; LENGTH: 890
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-26

Query Match          40.7%; Score 44; DB 6; Length 890;
Best Local Similarity 35.0%; Pred. No. 11;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy      2 PROQIILAGVFESALVDK 21
Db      375 PLQQTILFGLCNSSSIIEK 394

RESULT 3
US-09-978-360A-472
; Sequence 472, Application US/09978360A
; Publication No. US2006009633A9
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.USA.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 472
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-360A-472

Query Match          38.9%; Score 42; DB 5; Length 87;
Best Local Similarity 40.9%; Pred. No. 1.4;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy      2 PROQIILAGVFESALVDK 23
Db      61 PEQAFYWGPIEAVAKDKLA 82

RESULT 4
US-11-144-236-6
; Sequence 6, Application US/11144236
; Publication No. US20050288219A1
; GENERAL INFORMATION:
; APPLICANT: Nestec SA
; TITLE OF INVENTION: Osteoprotegerin in Milk

; FILE REFERENCE: 88265-6852
; CURRENT APPLICATION NUMBER: US/11/144,236
; CURRENT FILING DATE: 2005-06-02
; PRIOR APPLICATION NUMBER: US/10/676,358
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: WO 2002 EP 02912
; PRIOR FILING DATE: 2003-03-15
; PRIOR APPLICATION NUMBER: EP 20010108414
; PRIOR FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent version 3.1
; SEQ ID NO 6
; LENGTH: 537
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: protein sequence including mature OPG
US-11-144-236-6

Query Match          38.4%; Score 41.5; DB 7; Length 537;
Best Local Similarity 43.5%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

Qy      1 DPROQIILAGVFESALVDK 23
Db      63 DEKQFI---VFPDSATVDQIA 82

RESULT 5
US-09-978-360A-766
; Sequence 766, Application US/09978360A
; Publication No. US2006009633A9
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56.USA.CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 766
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -13...-1
US-09-978-360A-766
```


RESULT 10
US-11-010-795-18
; Sequence 18, Application US/11010795/
; Publication No. US30060005271A1
; GENERAL INFORMATION:
; APPLICANT: TUMER, NITGUN E.
; APPLICANT: DI, RONG

```

; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING L3 DELTA PROTEINS ARE
; FILE REFERENCE: RESISTANT TO TRICHOHECENE FUNGAL TOXINS
; CURRENT APPLICATION NUMBER: US/11/010,795
; PRIOR FILING DATE: 2004-12-13
; PRIOR APPLICATION NUMBER: 60/529,348
; PRIOR FILING DATE: 2003-12-12
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 18
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-11-010-795-18

Query Match      35.2% Score 38; DB 7; Length 387;
Best Local Similarity 30.0%; Pred. No. 42;
Matches 6; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY      4 QOLIALGAVFESALVDKLA 23
DB      200 EKTVAVDSEFQENEMIDA 219

RESULT 11
US-11-133-360-20
; Sequence 20, Application US/11133360
; Publication No. US20050287172A1
; GENERAL INFORMATION:
; APPLICANT: Medimmune Vaccines, Inc et al.
; TITLE OF INVENTION: INFLUENZA HEMAGGLUTININ AND NEURAMINIDASE VARIANTS
; FILE REFERENCE: FL260
; CURRENT APPLICATION NUMBER: US/11/133,360
; PRIOR FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: 60/574,553
; PRIOR FILING DATE: 2004-05-25
; PRIOR APPLICATION NUMBER: 60/657,554
; PRIOR FILING DATE: 2005-02-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Influenza virus
US-11-133-360-20

Query Match      35.2% Score 38; DB 7; Length 469;
Best Local Similarity 36.8%; Pred. No. 53;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      1 DPROOLALGAVFESALV 19
DB      2 NPNQKIAGSVSLTIAT 20

RESULT 12
US-11-133-346-20
; Sequence 20, Application US/11133346
; Publication No. US20060008473A1
; GENERAL INFORMATION:
; APPLICANT: Medimmune Vaccines, Inc et al.
; TITLE OF INVENTION: INFLUENZA HEMAGGLUTININ AND NEURAMINIDASE VARIANTS
; FILE REFERENCE: FL260
; CURRENT APPLICATION NUMBER: US/11/133,346
; PRIOR FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: 60/574,553
; PRIOR FILING DATE: 2004-05-25
; PRIOR APPLICATION NUMBER: 60/657,554
; PRIOR FILING DATE: 2005-02-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 469
```

```

; TYPE: PRT
; ORGANISM: Influenza virus
US-11-133-346-20

Query Match      35.2% Score 38; DB 7; Length 469;
Best Local Similarity 36.8%; Pred. No. 53;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY      1 DPROOLALGAVFESALV 19
DB      2 NPNQKIAGSVSLTIAT 20

RESULT 13
US-11-069-642-21
; Sequence 21, Application US/11069642
; Publication No. US20050260626A1
; GENERAL INFORMATION:
; APPLICANT: LORENS, JAMES B.
; APPLICANT: PRAY, TODD R.
; APPLICANT: KINSELLA, TODD M.
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR
; FILE REFERENCE: RIGL-022CIP3
; CURRENT APPLICATION NUMBER: US/11/069,642
; PRIOR FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 10/232,758
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 10/422,536
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/187,130
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii OT3
US-11-069-642-21

Query Match      35.2% Score 38; DB 7; Length 527;
Best Local Similarity 42.1%; Pred. No. 62;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY      5 QLIAGAVFESALVDKLA 23
DB      8 KVIANGQLFELRELVEKIS 26

RESULT 14
US-11-143-980-49
; Sequence 49, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: He, Min
; APPLICANT: Hucul, John
; APPLICANT: Haltli, Bradley A.
; APPLICANT: Wagenaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Summers, Mia
; APPLICANT: Kulowski, Kerry
; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; PRIOR FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
```

/ PRIOR FILING DATE: 2004-06-03
 / NUMBER OF SEQ ID NOS: 72
 / SOFTWARE: PatentIn version 3.3
 / SEQ ID NO 49
 / LENGTH: 7968
 / TYPE: PRT
 / ORGANISM: Streptomyces sp.
 US-11-143-980-49

Query Match 34.7%; Score 37.5; DB 7; Length 7968;
 Best Local Similarity 45.5%; Pred. No. 2.2e+03;
 Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 1 DPROQIALGAVF---ESALV 19
 DB 7225 DPRSALAAAGLVLADETTVMV 7246

RESULT 15
 US-10-516-768-18
 / Sequence 18, Application US/10516768
 / Publication No. US20050256302A1
 / GENERAL INFORMATION:
 / APPLICANT: MINAMINO, NAOTO
 / APPLICANT: KATAFUCHI, TAKESHI
 / TITLE OF INVENTION: NOVEL PEPTIDES HAVING CAMP PRODUCING ACTIVITY
 / FILE REFERENCE: 62273(71526)
 / CURRENT APPLICATION NUMBER: US/10/516,768
 / CURRENT FILING DATE: 2004-12-03
 / PRIOR APPLICATION NUMBER: PCT/JP03/06641
 / PRIOR FILING DATE: 2003-05-28
 / PRIOR APPLICATION NUMBER: JP 2002-162797
 / PRIOR FILING DATE: 2002-06-04
 / NUMBER OF SEQ ID NOS: 52
 / SOFTWARE: PatentIn Ver. 3.3
 / SEQ ID NO 18
 / LENGTH: 125
 / TYPE: PRT
 / ORGANISM: Sus sp.
 US-10-516-768-18

Query Match 34.3%; Score 37; DB 6; Length 125;
 Best Local Similarity 46.7%; Pred. No. 15;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 IALGAVFESALVDK 21
 DB 30 MALGSSPDSATLTER 44

Search completed: January 23, 2006, 11:53:36
 Job time : 6.1 secs

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838A-6
Perfect score: 108
Sequence: 1 DPROQIALGAVFESALVDKLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Genesegp21:*
2: genesegp1908:*
3: genesegp2000:*
4: genesegp2001:*
5: genesegp2002:*
6: genesegp2003a:*
7: genesegp2003b:*
8: genesegp2004a:*
9: genesegp2005a:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	23	6	ADBI6891 Pseudomon
2	108	100.0	23	9	AEBO8223 Pseudomon
3	108	100.0	206	9	AEBO8266 Pseudomon
4	108	100.0	495	7	ABO82704 Pseudomon
5	48	44.4	606	9	ADW70865 A. thalita
6	47	43.5	381	2	AAV06368 Streptomy
7	47	43.5	381	3	AAAI4881 Rhodocet
8	47	43.5	381	5	AAU77589 S. livida
9	47	43.5	381	5	AAU77433 Streptomy
10	47	43.5	429	3	AAV64346 Aay84346 Amino aci
11	47	43.5	878	4	AAV73265 Yeast tte
12	47	43.5	888	8	ADU49053 OLI-asso
13	47	43.5	3129	3	AAAB07427 Amino aci
14	47	43.5	3129	9	ADZ67207 Fusaarium
15	46	42.6	107	5	AAAM5054 Bartonell
16	46	42.6	345	8	ADT91429 Arabidops
17	46	42.6	357	6	ADT57742 plant pol
18	46	42.6	441	6	ADAI3778 Actinobda
19	46	42.6	492	6	ADVI6514 E. faecal
20	46	42.6	493	7	ADH87662 Enterococ
21	46	42.6	632	7	ADG30712 Xanthomon
22	45.5	42.1	334	6	ABU9254 Protein e
23	45.5	42.1	334	7	ADP28252 Bacterial
24	45.5	42.1	334	7	ADG32461 Pasteurel

25	45	41.7	173	4	ABG09047
26	45	41.7	355	5	ABG91507
27	44	40.7	77	4	AAU23766
28	44	40.7	247	4	AAU63428
29	44	40.7	247	6	ABM59947
30	44	40.7	304	9	ABBA9954
31	44	40.7	352	4	ABG25486
32	44	40.7	535	6	ABU26050
33	44	40.7	535	9	ABM95939
34	44	40.7	585	8	ADJ34656
35	44	40.7	593	8	ADG73807
36	44	40.7	618	8	ADG30171
37	44	40.7	878	4	ABG25491
38	44	40.7	890	8	ADG32056
39	43.5	40.3	361	6	ABM68894
40	43.5	40.3	438	4	ABB69582
41	43	39.8	102	2	AAW87512
42	43	39.8	124	2	AAW78909
43	43	39.8	129	2	AAW45188
44	43	39.8	129	2	AAW06813
45	43	39.8	129	2	AAW78905

ALIGNMENTS

RESULT 1
ADBI6891
ID ADBI6891 standard; peptide; 23 AA.
XX

ADBI6891;

20-NOV-2003 (first entry)

Pseudomonas aeruginosa PA2627 penetrating peptide 6.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;
infertility; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
osteopathic; cytostatic; nootropic.

Pseudomonas aeruginosa.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-IB000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating
or preventing e.g. endocrine disorders.

Claim 2; Page 14; 60pp; English.

This invention relates to a novel peptide sequences capable of
translocating across a biological barrier. Furthermore, it refers to
methods that use these peptides to facilitate penetration of a
biologically active effector molecule such as a drug or other therapeutic
agent across biological barriers e.g. epithelial or endothelial cells
sealed by tight junctions. This peptide is derived from a bacterial
toxin, an integral membrane or extracellular protein and can comprise an
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
or enzyme. The effector molecule, however, can comprise for example
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
CC cyostatic or neurotropic activities. This peptide is from PA2627 of
CC Pseudomonas aeruginosa and is penetrating peptide 6 of the invention.
XX
SQ Sequence 23 AA;
Query Match 100.0%; Score 108; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 1,7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DPROQLALGAVFESALVDKLA 23
Db 1 DPROQLALGAVFESALVDKLA 23
RESULT 2
AEB08223 standard; peptide; 23 AA.
AC AEB08223;
DT 25-AUG-2005 (first entry)
XX
XX Pseudomonas aeruginosa PA2627 penetrating peptide 6, SEQ ID NO: 6.
XX
KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
KW antidiabetic; endocrine disease; gastrointestinal disease;
KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
KW neurotropic; neurological disease; parkinsons disease; antiparkinsonian;
KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
KW anticonvulsant; genetic disorder; cardiovascular disease;
KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
KW genitourinary disease; hematological disease; antianemic; anemia;
KW autoimmune disease; immunosuppressive; immune deficiency;
KW immunostimulant; infectious disease; antimicrobial; infection;
KW erectile dysfunction; andrology; major depressive disorder;
KW antidepressant; psychiatric disorder; pain; analgesic;
KW bacterial infection; antibacterial; viral infection; virocidic;
KW fungal infection; fungicide; parasitic infection; antiparasitic;
KW renal failure; antifertility; antirheumatic; cyostatic;
KW antiinflammatory; hepatotropic; hepatitis B virus infection.
XX
OS Pseudomonas aeruginosa.
PN US2005136103-A1.
XX
PD 23-JUN-2005.
XX
PF 16-SEP-2004; 2004US-00942300.
XX
PR 17-SEP-2003; 2003US-00664989.
PR 17-SEP-2003; 2003US-00665184.
PR 17-SEP-2003; 2003US-0503615P.
XX
PA (BENS/) BEN-SANSON S A.
PA (COHE/) COHEN E.
XX
PI Ben-Sanson SA, Cohen E;
XX
XX WPI; 2005-444089/45.
PT Composition used for translocating effectors across barrier such as
PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
XX
XX Claim 46; SEQ ID NO 6; 59pp; English.
PS
XX The present invention relates to a pharmaceutical composition of
CC penetrating peptides for transcellular delivery of effector. The
CC invention comprises the effector sequentially coupled with a counter ion
CC and at least one hydrophobic agent, where the effector is selectively
CC encapsulated into a complex. The invention is useful for translocating
CC effectors across a biological barrier such as epithelial cells and
CC endothelial cells during treatment and prevention of disease or
CC pathological conditions (including endocrine disorders, diabetes,
CC infertility, hormone deficiencies, osteoporosis, ophthalmological
CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
CC coagulable states, coronary disease, cerebrovascular events, metabolic
CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC hematologic disorders, anemia of different entities, immunologic and
CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
CC infectious diseases, viral infections, bacterial infections, fungal
CC infections, parasitic infections, neoplastic diseases, multi-factorial
CC disorders, impotence, chronic pain, depression, different fibrosis states
CC and short stature) and for mucosal vaccination against anthrax and
CC hepatitis B. The present sequence is the Pseudomonas aeruginosa PA2627
CC penetrating peptide. This sequence is used in the effective translocation
CC of aminoglycoside antibiotics and antifungal agents across an epithelial
CC barrier.
XX
SQ Sequence 23 AA;
Query Match 100.0%; Score 108; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 1,7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DPROQLALGAVFESALVDKLA 23
Db 1 DPROQLALGAVFESALVDKLA 23
RESULT 3
AEB08266 standard; protein; 206 AA.
ID AEB08266;
AC AEB08266;
DT 25-AUG-2005 (first entry)
XX
XX Pseudomonas aeruginosa protein, SEQ ID NO: 64.
XX
KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
KW antidiabetic; endocrine disease; gastrointestinal disease;
KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
KW neurotropic; neurological disease; parkinsons disease; antiparkinsonian;
KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
KW anticonvulsant; genetic disorder; cardiovascular disease;
KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
KW genitourinary disease; hematological disease; antianemic; anemia;
KW autoimmune disease; immunosuppressive; immune deficiency;
KW immunostimulant; infectious disease; antimicrobial; infection;
KW erectile dysfunction; andrology; major depressive disorder;
KW antidepressant; psychiatric disorder; pain; analgesic;
KW bacterial infection; antibacterial; viral infection; virocidic;
KW fungal infection; fungicide; parasitic infection; antiparasitic;
KW renal failure; antifertility; antirheumatic; cyostatic;
KW antiinflammatory; hepatotropic; hepatitis B virus infection.
XX
OS Pseudomonas aeruginosa.

PN US2005136103-A1.

XX 23-JUN-2005.

XX 16-SEP-2004; 2004US-00942300.

XX 17-SEP-2003; 2003US-00664989.

XX 17-SEP-2003; 2003US-00665184.

XX 17-SEP-2003; 2003US-0503615P.

XX (BENS/) BEN-SAMSON S. A.

XX (COHE/) COHEN E.

XX Ben-Samson SA, Cohen E;

XX WPI, 2005-444089/45.

XX Composition used for translocating effectors across barrier such as

XX epithelial cells during treatment of e.g. endocrine disorders comprises

XX effector sequentially coupled with counter ion and hydrophobic agent.

XX Disclosure; SEQ ID NO 64; 59pp; English.

XX The present invention relates to a pharmaceutical composition of

XX penetrating peptides for transmembrane delivery of effector. The

XX invention comprises the effector sequentially coupled with a counter ion

XX and at least one hydrophobic agent, where the effector is selectively

XX encapsulated into a complex. The invention is useful for translocating

XX effectors across a biological barrier such as epithelial cells and

XX endothelial cells during treatment and prevention of disease or

XX pathological conditions (including endocrine disorders, diabetes,

XX infertility, hormone deficiencies, osteoporosis, ophthalmological

XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,

XX Parkinson's disease, multiple sclerosis, Huntington's disease,

XX cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-

XX coagulable states, coronary disease, cerebrovascular events, metabolic

XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,

XX hematological disorders, anemia of different entities, immunologic and

XX rheumatologic disorders, autoimmune diseases, immune deficiencies,

XX infectious diseases, viral infections, bacterial infections, fungal

XX infections, parasitic infections, neoplastic diseases, multi-factorial

XX disorders, impotence, chronic pain, depression, different fibrosis states

XX and short stature) and for mucosal vaccination against anthrax and

XX hepatitis B. The present sequence is the Pseudomonas aeruginosa protein

XX containing penetrating peptide at N-terminal end.

XX SQ Sequence 206 AA;

XX Query Match 100.0%; Score 108; DB 9; Length 206;

XX Best Local Similarity 100.0%; Pred. No. 2.5e-10;

XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 DPROQIALGAVFESALVDKLA 23

XX DB 3 DPROQIALGAVFESALVDKLA 25

XX RESULT 4

XX ID ABO82704 standard; protein; 495 AA.

XX AC ABO82704;

XX DT 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polypeptide #14879.

XX KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX OS Pseudomonas aeruginosa.

XX EN US6551995-B1.

PD 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nollig J, Deloughery C, Bush D;

XX WPI, 2003-615309/58.

XX N-PSDB; ABD16275.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

XX useful as molecular targets for diagnostics, prophylaxis and treatment of

XX pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 31450; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the

XX polynucleotides encoding them. The sequences are useful in diagnosis and

XX therapy of pathological conditions, as molecular targets for diagnostics,

XX prophylaxis and treatment of pathological conditions resulting from a

XX bacterial infection, for evaluating a compound, such as a polypeptide,

XX for the ability to bind a P. aeruginosa nucleic acid, as components of

XX effective antibacterial targets, as targets for antibacterial drugs,

XX including anti-P. aeruginosa drugs, as templates for recombinant

XX production of P. aeruginosa-derived peptides or polypeptides, as target

XX components for diagnosis and/or treatment of P. aeruginosa-caused

XX infection, and in detection of P. aeruginosa sequences or other sequences

XX of Pseudomonas species using bioclip technology. Sequences ABO67826-

XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

XX sequence data for this patent did not form part of the printed

XX specification but was obtained in electronic format from USPTO at

XX seqdata.uspto.gov/sequence.html

XX SQ Sequence 495 AA;

XX Query Match 100.0%; Score 108; DB 7; Length 495;

XX Best Local Similarity 100.0%; Pred. No. 7.5e-10;

XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 DPROQIALGAVFESALVDKLA 23

XX DB 292 DPROQIALGAVFESALVDKLA 314

XX RESULT 5

XX ADW70865 standard; protein; 606 AA.

XX AC ADW70865;

XX DT 21-APR-2005 (first entry)

XX DE A. thaliana CRR4 protein, seq id 2.

XX KM Plant engineering; chloroplast; CRR4; NDH; NADPH dehydrogenase.

XX OS Arabidopsis thaliana.

XX OS US2005026280-A1.

XX PD 03-FEB-2005.

XX 22-DEC-2003; 2003US-00740618.

XX 18-JUL-2003; 2003JP-00199098.

XX (TASAKI) TASAKI M.

XX (SHIKI) SHIKI M. T.

Query Match	43.5%	Score 47;	DB 3;	Length 381;
Best Local Similarity	57.1%	Pred. No. 31;		
Matches 12;	Conservative 1;	Mismatches 8;	Indels 0;	Gaps 0;

```

QY      2 PRQQLTALGAVFESALVDKL  22
          ||| ||||| : ||| |
DB      11 PRGLLAALGAVLAPALVSSL  31

```

RESULT 8
AAU77589
ID AAU77589 standard; protein: 381 AA.

AC AAU77589;

DT 05-JUN-2002 (first entry)

DB 8. 11vdans EGII-1like cellulase.

KM Eggr, cellulase; endoglucanase II; detergent; cellulose treatment;
 KM stonewashing; indigo dyed denim; feed additive; wood pulp treatment;
 KM blowabs reduction; laundry; dish detergent; milling; depilling;
 KM softening; surface fibre removal; anti-feyling.

Streptomyces lividans.

PN WO200212466-A2.

PD 14-FEB-2002.

31-JUL-2001; 2001WO-US023991.

04-AUG-2000; 2000US-00633085.

PA (GEMV) GENENCOR INT INC.

PI Day AG, Gualfetti P, Mitchinson C, Shaw A;

WPI; 2002-241752/29.

PT Novel variant of endoglucanase III or endoglucanase III-like cellulase for treating cellulose containing textile, has performance sensitive residues replaced to residue having modified stability.

Example 1; Fig 3; 47pp; English.

The invention relates to a variant of endoglucanase III (EGIII) or EGIII-like cellulase comprising a substitution or deletion at a position corresponding to one or more of residues W7, G31, A35, T145, Y147, Q162 and/or Y168 in EGIII from *Trichoderma reesei*. Also included are a DNA encoding the variant, a vector comprising the DNA, a host cell transformed with the vector and a detergent composition comprising a surfactant and the variant. The variant is useful in the treatment of a cellulose containing textile, stone-washing or indigo dyed denim or as a feed additive or in the treatment of wood pulp, in reduction of biomass to glucose. The detergent composition is useful as the main component of a laundry or dish detergent and is further useful as pre-wash composition, pre-soak composition or for cleaning during the regular wash or clean cycle. The variant increases value of animal feed, improves the drainability of food pulp, enhances food products and reduces fibre in grain during grain wet (or dry) milling process. Further cellulase improves the feel e.g. smoothness and/or appearance e.g. removing pills and fibrils which tend to reduce the sharpness in appearance of a fabric, of cellulose containing fabric, and imparts desirable effects such as depilling, softening, anti-pilling, surface fiber removal, anti-greying and cleaning. The present sequence represents an EGIII-like cellulase with homology to that of the T. reesei protein, encoded by a gene isolated by the primers appearing as ABK11339-ABK11349

Sequence 381 AA;
SQ

Query Match 43.5%; Score 47; DB 5; Length 381;

Best Local Similarity 57.1%; Pred. No. 31;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

```

QY      2 PRQQLIALGAVFESALVDKL 22
        ||| ||||| : ||| |
DB      11 PRGLLAALGAVLAAAFALVSSL 31

```

RESULT 9
AAU77433
ID AAU77433 standard; protein; 381 AA.

AAU77433;

DT 29-AUG-2003 (revised)

[illegible][illegible]

cellulose containing textile; enzyme.

OS Streptomyces lividans; CelB.

PN WO200212464-A2

PD 14-FEB-2002.

31-JUL-2001; 2001WO-US023989.

PR 04-AUG-2000; 2000US-00632426.

PA (GEMV) GENENCOR INT INC.

PI Mitchinson C, Ropp TH, Swanson BA;

DR WPI; 2002-241750/29.

AA Novel endoglucanase II (EGII)-like cellulase variant comprising
PT substitution/deletion at positions corresponding to specific residues in
PT EGII from *Trichoderma reesei*, useful for treating cellulose containing
PT textile.

Example 1; Fig 3; 41pp; English.

XX The present invention relates to novel endoglucanase III (BGIII)-like
CC cellulase variants which comprise a substitution or deletion at a
CC position corresponding to one or more of residues M79, M154 and/or M118
CC in mature BGIII from the fungus, *Trichoderma reesei*. The variants are
CC useful in the treatment of a cellulose containing textile. By
CC substituting other amino acids for the native methionines at positions
CC 79, 118 and 154 (sites where oxidation of the enzyme takes place) in
CC BGIII from *T. reesei*, oxidatively more stable enzymes are obtained. The
CC present sequence represents *Streptomyces lividans* CabI BGIII-like
CC cellulase. (Updated on 29-AUG-2003 to standardise OS field)

AA Sequence 381 AA;
SQ

Query Match	43.5%	Score 47;	DB 5;	Length 381;
Best Local Similarity	57.1%;	Pred. No. 31;		
Matches 12; Conservative	1;	Mismatches 8;	Indels 0;	Gaps 0;

```

QY      2 PRQQLIAGAVFESAAVDKL 22
      ||| | | | | | : | | | |
Db      11 PRGLIAGAVTAAFAVSSL 31

```

RESULT 10
AAY84346

AA
AC
AA84346

XXV

DT 12-JUN-2000 (first entry)
XX Amino acid sequence of an endoglucanase III (EGIII)-like cellulase.
DE
XX Endoglucanase III; EGIII; EGIII-like cellulase; surfactant stability;
KW cellulase; textile processing; textile cleaning; stonewashing;
KW indigo dyed denim; cellulose containing fabric; fabric smoothness;
KW pill removal; fibril removal; cotton; cellulosic fibre; dyeing; detergent;
KW animal feed; wood pulp; paper; grain; biomass reduction; glucose.
XX
OS Streptomyces lividans.
XX
PN MO200014208-A1.
XX
PD 16-MAR-2000.
XX
XX 24-AUG-1999; 99WO-US019154.
XX
PR 03-SEP-1998; 98US-00146729.
XX
PA (GENV) GENENCOR INT INC.
XX
XX Fowler T;
XX
XX WPI, 2000-271052/23.
XX
XX Novel variant endoglucanase III-like cellulases with improved surfactant
PT stability and resistance to temperature stress, useful for textile
PT processing or cleaning, treating wood pulp, food and grain, and reducing
PT biomass to glucose.
XX
XX Disclosure; Page 65-66; 73pp; English.
XX
XX The present sequence represents an endoglucanase III (EGIII)-like
CC cellulase. The cellulase has homology to the Trichoderma reesei EGIII
CC protein. The variant cellulases have improved temperature stability, and
CC improved surfactant stability. The variant cellulases and compositions
CC containing them are used in textile processing or cleaning, e.g. in
CC stonewashing of indigo dyed denim, and modifying the texture, feel or
CC appearance of cellulose containing fabrics (e.g. improving fabric
CC smoothness or removing pills and fibrils). The compositions may also be
CC used for the removal of immature or dead cotton from cellulosic fibres or
CC fabric, which can cause uneven dyeing. The cellulase may also be used in a
CC detergent composition for washing laundry and dishes and in the treatment
CC of animal feed, wood pulp, paper, non-animal foods and grains. The
CC enzymes may also be used in the reduction of biomass to glucose.
CC
XX
SQ Sequence 429 AA;
Query Match 43.5%; Score 47; DB 3; Length 429;
Best Local Similarity 57.1%; Pred. No. 35; Mismatches 8; Indels 0; Gaps 0;
Matches 12; Conservative 1;
QY 2 PROQIAGAVFESALVDKL 22
DB 11 PRGLAALGAVLAARLVSSL 31

RESULT 11
AAB73265
ID AAB73265 standard; protein; 878 AA.
XX
XX AAB73265;
XX
DT 17-MAY-2001 (first entry)
XX
DE Yeast trehalose-6-phosphate phosphatase TPP.
XX
KW Yeast; TPP; antiparasitic; fungicide; bactericide; protozoacide;
KW insecticide; trehalose-6-phosphate phosphatase.
XX
OS Candida albicans.
XX

PN MO200116357-A2.
XX
XX 08-MAR-2001.
PD
XX 29-AUG-2000; 2000WO-EP008410.
XX
XX 30-AUG-1999; 99EP-00202805.
PR 27-JUN-2000; 2000EP-00870145.
XX
XX (LEUV-) LEUVEN RES & DEV.
XX
XX Thevelein J, Van Dijk P;
PI
DR WPI; 2001-235118/24.
XX
XX
XX New inhibitors and screening assay for inhibitors or suppressors of sugar
PT alcohol phosphatases or sugar phosphatases.
PT
XX
XX Disclosure; Fig 8; 106pp; English.
PS
XX The present invention relates to a screening assay for inhibitors or
CC suppressors of sugar alcohol phosphatases or sugar phosphatases using
CC enzymes found in fungi, bacteria, insects, nematodes, mites or protozoa.
CC The inhibitors identified are useful for treating or preventing
CC parasitic, particularly fungal, bacterial or protozoal infections, or a
CC nematode, insect, worm or mite infestation, in a human, animal or plant.
CC The present sequence is one such sugar phosphatase: yeast trehalose-6-
CC phosphate phosphatase (TPP)
XX
XX
SQ Sequence 878 AA;
Query Match 43.5%; Score 47; DB 4; Length 878;
Best Local Similarity 46.2%; Pred. No. 85;
Matches 12; Conservative 3; Mismatches 5; Indels 6; Gaps 1;
QY 1 DPRQIAGG-----AVFESALVD 20
DB 817 EPRQLETLGLAGLVISFESAGTVD 842

RESULT 12
ADJ49053
ID ADJ49053 standard; protein; 888 AA.
XX
XX ADJ49053;
XX
DT 06-MAY-2004 (first entry)
XX
DE Oil-associated gene related protein #553.
XX
KW oil-associated gene; transgenic; enhanced seed oil; vegetable oil.
XX
OS Unidentified.
XX
XX US2004025202-A1.
PN
XX
PD 05-FEB-2004.
XX
XX 14-MAR-2003; 2003US-00389566.
PF
XX 15-MAR-2002; 2002US-0365301P.
PR 26-JUN-2002; 2002US-0391786P.
PR 26-JUN-2002; 2002US-0392018P.
XX
XX (LAUR/) LAURIE C C.
PA (RAVA/) RAVANELLO M.
PA (SAVA/) SAVAGE T.
PA (LEDE/) LEDEAUX J R.
PA (ROGE/) ROGERS J A.
XX
XX Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;
PI
XX WPI; 2004-142683/14.
XX
XX

```

XX Novel recombinant DNA construct comprising a promoter functional in
XX plants operably linked to an oil-associated gene for producing transgenic
XX plant seed.
XX
XX Example 3; SEQ ID NO 1057; 22pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in plants operably linked to an oil-associated gene.
XX The construct is useful for transgenic plant seed which has in its genome
XX the construct, that is functional in the plant to transcribe the oil-
XX associated gene. The transgenic plant seed grows into a plant having
XX enhanced seed oil as compared to wild type. The construct is useful for
XX producing hybrid maize seed. The transgenic plant seed is useful for
XX producing vegetable oil. The present sequence represents the amino acid
XX sequence of an oil-associated gene related protein.
XX
XX Sequence 888 AA;
XX
XX Query Match 43.5%; Score 47; DB 8; Length 888;
XX Best Local Similarity 46.2%; Pred. No. 86;
XX Matches 12; Conservative 3; Mismatches 5; Indels 6; Gaps 1
XX
XX 1 DPROQLIAG-----AVFESALVD 20
XX :||| ||| :||| |||
XX 817 EPROVLETLGLAGLVSTFESAGTYD 842
XX
XX RESULT 13
XX AAB07427
XX ID AAB07427 standard; protein; 3129 AA.
XX
XX AAB07427;
XX
XX 20-OCT-2000 (first entry)
XX
XX Amino acid sequence of a cyclohexadepsipeptide synthetase.
XX
XX Cyclohexadepsipeptide synthetase; filamentous fungal cell;
XX cyclohexadepsipeptide; antibiotic.
XX
XX Fusarium venenatum.
XX
XX WO200042203-A2.
XX
XX 20-JUL-2000.
XX
XX 13-JAN-2000; 2000MO-US000913.
XX
XX 13-JAN-1999; 99US-00229862.
XX
XX (NOVO ) NOVO NORDISK BIOTECH INC.
XX
XX Berka RM, Rey MM, Yoder WT;
XX
XX WI; 2000-482833/42.
XX
XX N-PSDB; AAA58762.
XX
XX Producing a heterologous polypeptide for production of antibiotics
XX comprises cultivating a mutant of a parent filamentous fungal cell
XX comprising a nucleic acid sequence encoding cyclohexadepsipeptide.
XX
XX Claim 30; Fig 1A-I; 76pp; English.
XX
XX The present sequence represents a cyclohexadepsipeptide synthetase
XX polypeptide. The specification describes a method for producing a
XX heterologous polypeptide. The method comprises cultivating a mutant of a
XX parent filamentous fungal cell, which produces less cyclohexadepsipeptide
XX than the parent filamentous fungal cell when cultured under the same
XX conditions. The method is used for the production of biologically active
XX compounds e.g. antibiotics
XX
XX Sequence 3129 AA;

```

Query Match	43.5%	Score 47	DB 3	Length 3129
Best Local Similarity	43.5%	Pred. No. 4e+02		
Matches	10	Conservative 3	Mismatches 10	Indels 0
				Gaps 0

DB	1	DPROOGLNAGVFESALVNDLA 23
1997	DPEOOLVGIGWGLVVTGDLA	1919

RESULT 14
AD267207
ID AD267207 standard; protein: 3129 AA.
AC AD267207;
XX cyclohexadepsipeptide-deficient filamentous fungal mutant cells.
XX 28-JUL-2005 (first entry)
DT
DE Fusarium venenatum cyclohexadepsipeptide synthase.
XX cloning; mutant; protein secretion; cyclohexadepsipeptide synthase;
KW enzyme.
XX Fusarium venenatum.
OS
XX US6893839-B1.
PN
XX 17-MAY-2005.
PD
XX 13-JAN-2000; 2000US-00482788.
PF
XX 13-JAN-1999; 99US-00229862.
PR
XX (NOVO) NOVOZYMES BIOTECH INC.
PA
XX Berka RM, Rey KM, Yoder WT;
PI
XX WPI: 2005-342680/35.
DR
XX N-PSDB; AD267206.
PT
XX Producing a secreted heterologous polypeptide, useful for producing
PT cyclohexadepsipeptide-deficient filamentous fungal mutant cells,
PT comprises cultivating a mutant cell of a parent Fusarium venenatum.
XX
XX Claim 5; SEQ ID NO 2; 40pp; English.
PS
XX The invention relates to a method of producing a secreted heterologous
XX polypeptide comprising cultivating a mutant cell of a parent Fusarium
CC venenatum cell under conditions for the production of the secreted
CC heterologous polypeptide. The method comprises: (a) cultivating a mutant
CC cell of a parent F. venenatum cell under conditions for the production of
CC the secreted heterologous polypeptide, where: (i) the mutant cell
CC comprises a first nucleic acid encoding the secreted heterologous
CC polypeptide; and (ii) the mutant cell comprises a second nucleic acid
CC which comprises a disruption or a deletion in a cyclohexadepsipeptide
CC synthetase gene, where the mutant cell produces less
CC cyclohexadepsipeptide than the parent F. venenatum cell when cultured
CC under the same conditions as a result of the disruption or the deletion
CC in the cyclohexadepsipeptide synthetase gene, where the
CC cyclohexadepsipeptide synthetase gene encodes a cyclohexadepsipeptide
CC synthetase having an amino acid sequence which has at least 95% identity
CC with a sequence of 3129 amino acids, or a cyclohexadepsipeptide
CC synthetase which is encoded by a nucleic acid which hybridizes under at
CC least high stringency conditions with (i) the nucleic acid of 11212 base
CC pairs (bp), (ii) the corresponding cDNA; or (iii) a complete
CC complementary strand of (i) or (ii), and where high stringency conditions
CC are defined as prehybridization and hybridization at 45 deg C in 5xSSPE,
CC 0.3% SDS, 200 g/ml sheared and denatured salmon sperm DNA, and 50%
CC formamide and washing three times each for 15 minutes using 2xSSC, 0.2%
CC SDS at 55 deg C; and (b) isolating the secreted heterologous polypeptide
CC from the cultivation medium. The method is useful for producing a
CC secreted heterologous polypeptide. This sequence corresponds to the
CC Fusarium venenatum cyclohexadepsipeptide synthetase of the invention.

XX Sequence 3129 AA;

Job time : 72.2 secs

Query Match 43.5%; Score 47; DB 9; Length 3129;
 Best Local Similarity 43.5%; Pred. No. 4e+02;
 Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 DPROQLALGAVFESALVDKLA 23
 |||||:|:|
 Db 1897 DPEQDLVGIGWGLVVTGDGLA 1919

RESULT 15

AAMS0554
 ID AAMS0554 standard; protein; 107 AA.

AC AAMS0554;

DT 18-MAR-2002 (first entry)

XX Bartonella henselae VirB2 virulence protein.

KW VirB2; virulence; cat scratch disease; bacillary angiomatosis; infection;
 therapy; diagnosis; vaccine.

OS Bartonella henselae.

PN W0200192535-A1.

PD 06-DEC-2001.

PF 05-JUN-2000; 2000WO-US015465.

PR 01-JUN-2000; 2000US-0208799P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Padmalayam I, Masung R, Karem K, Baumstark B;

DR WPI; 2002-114352/15.

DR N-PSDB; AAI70976.

PT Novel nucleic acids of Bartonella henselae virB virulence operon and
 PT polypeptides encoded by the nucleic acids useful for diagnosing, treating
 PT Bartonella henselae infection and cat scratch disease.

PS Claim 33; Fig 1; 114pp; English.

XX The present sequence is that of virulence protein VirB2 of Bartonella
 CC henselae strain Houston-1, as deduced from an open reading frame in the
 CC newly isolated virB operon (see AAI70976). The protein was identified by
 CC its homology to the virB2 protein of Agrobacterium tumefaciens. It has a
 CC predicted mol.wt. of 10,900. Isolated VirB polypeptides of B. henselae
 CC are specifically claimed. They can be produced by recombinant expression
 CC in prokaryotic or eukaryotic host cells, and are used in claimed
 CC therapeutic compositions, including vaccines, against B. henselae
 CC infection. Antibodies raised against the VirB polypeptides are used in a
 CC claimed method of detecting current or previous B. henselae infection. B.
 CC henselae is the causative agent of cat scratch disease and bacillary
 CC angiomatosis

CC Sequence 107 AA;

Query Match 42.6%; Score 46; DB 5; Length 107;

Best Local Similarity 58.8%; Pred. No. 9.7;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 6 LIALGAVFESALVDKLA 22
 |||||:|:|

Db 87 LIGIGIVGASALVSKL 103

Search completed: January 23, 2006, 09:20:50

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-6

Perfect score: 108
Sequence: 1 DPROQIALGAVFESALVDKLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	206	2 D83317	conserved hypochet
2	44.4	44.4	605	2 B84889	hypothetical prote
3	47	43.5	594	2 D96015	probable tartonat
4	47	43.5	1494	2 T26452	hypothetical prote
5	46	42.6	364	2 T02364	hypothetical prote
6	45.5	42.1	930	2 AF3431	pyruvate, phosphat
7	45	41.7	222	2 T35942	probable ABC-type
8	45	41.7	260	2 AD1829	ATP-binding protei
9	44	40.7	241	2 H82435	conserved hypochet
10	44	40.7	327	2 T44255	probable amino aci
11	44	40.7	516	2 AB3438	amldophosphoribosy
12	44	40.7	593	1 JT0742	tartonnate-semiald
13	44	40.7	593	2 C85550	glyoxylate carbol
14	44	40.7	593	2 H90639	glyoxylate carbol
15	43.5	39.8	441	2 B84103	late competence pr
16	43	39.8	344	2 A72780	probable ribosomal
17	43	39.8	437	2 S75730	8-amino-7-oxonon
18	43	39.8	459	2 T00129	hypothetical prote
19	43	39.8	591	2 B83457	glyoxylate carbol
20	43	39.8	593	2 AH0566	tartonnate-semiald
21	43	39.8	680	2 A37890	NADPH-ferrihempro
22	43	39.8	724	2 S42868	serine/threonine p
23	43	39.8	3131	2 S39842	emulatin synthet
24	42.5	39.4	463	2 T51313	hypothetical prote
25	42	38.9	207	2 A87543	conserved hypochet
26	42	38.9	370	2 C83120	probable RND efflu
27	42	38.9	396	2 AH2859	MPS permease limpo
28	42	38.9	396	2 F97636	probable transport
29	42	38.9	472	2 A28701	H+-transporting tw

30	42	38.9	496	2 AD1211	B. subtilis yuek p
31	42	38.9	496	2 AD1567	conserved hypochet
32	42	38.9	528	1 FMR08	H+-transporting tw
33	42	38.9	529	2 A33370	H+-transporting tw
34	42	38.9	593	2 T32485	hypothetical prote
35	42	38.9	599	2 S65180	hypothetical prote
36	42	38.9	815	2 A64048	aspartate kinase (
37	42	38.9	896	2 S48761	trehalose-phosphat
38	42	38.4	454	1 A26955	alkaline serine pr
39	41.5	38.4	953	1 S56650	pyruvate, phosphat
40	41.5	38.4	953	2 S12894	pyruvate, phosphat
41	41.5	38.4	955	2 S56649	pyruvate, phosphat
42	41.5	38.4	956	1 S53297	pyruvate, phosphat
43	41.5	38.4	1048	2 T23764	hypothetical prote
44	41	38.0	210	2 AD2570	transposase ali855
45	41	38.0	303	2 G83093	UDP-3-O-acyl-N-ace

ALIGNMENTS

RESULT 1

conserved hypothetical protein PA2627 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence=revision15#seq=2000#text=change_09-Jul-2004

C/Accession: D83317

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim

Loiy, S.; Olson, M.W.

Nature 406:165-169 (2000)

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A/Reference number: A82950; WUID:2043737; PMID:10984043

A/Accession: D83317

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1206 <SNO>

A/Cross-references: UNIPROT:Q910L1; UNIPARC:UP100000C5774; GB:AE004691; GB:AE004091; NI

A/Experimental source: strain PA01

C/Genetics:

A/Gene: PA2627

Query Match 100.0%; Score 108; DB 2; Length 206;

Best Local Similarity 100.0%; Pred. No. 4.6e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 DPROQIALGAVFESALVDKLA 23

3 DPROQIALGAVFESALVDKLA 25

RESULT 2

B84889

hypothetical protein At2g45350 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence=revision 02-Feb-2001 #text=change 09-Jul-2004

C/Accession: B84889

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: B84889

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1606 <SNO>

A/Cross-references: UNIPROT:O22137; UNIPARC:UP100000A11B6; GB:AE002093; NID:g2583119; P

C/Genetics:

A/Gene: At2g45350

A/Map position: 2

C/Superfamily: Arabidopsis thaliana hypothetical protein At2g41080

submitted to the EMBL Data Library, January 1999
A/Reference number: Z21551

A/Accession: T35942

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-222 <SEE>

A/Cross-references: UNIPROT:Q9ZBF9; UNIPARC:UPI00000DA9E; EMBL:AL035206; PIDD:CAA22764.

A/Experimental source: strain A3(2)

A/Genetics:

A/Genes: SC0DB:SC9B5.22

Query Match 41.7%; Score 45; DB 2; Length 222;

Best Local Similarity 52.9%; Pred. No. 11;

Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 PROQIALGAVFESALVDK 18

DB 74 PROQIALGAVFESALVDK 90

RESULT 8

AD1829
ATP-binding protein of ABC transporter at0180 [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120

C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C/Accession: AD1829

R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*

A/Reference number: AB1807; MUID:21595285; PMID:11759840

A/Accession: AD1829

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-260 <KUR>

A/Cross-references: UNIPROT:Q8Z0B8; UNIPARC:UPI00000CDC68; GB:BA000019; PIDD:BA077704.1;

A/Experimental source: strain PCC 7120

C/Genetics:

A/Genes: at0180

C/Superfamily: inner membrane protein malK; ATP-binding cassette homology

Query Match 41.7%; Score 45; DB 2; Length 260;

Best Local Similarity 47.4%; Pred. No. 13;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 QOLIALGAVFESALVDK 22

DB 77 QDPVIGWVFOQALPDSL 95

RESULT 9

H82435
conserved hypothetical protein VCA0634 [imported] - *Vibrio cholerae* (strain N16961 serog

C/Species: *Vibrio cholerae*

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C/Accession: H82435

R/Heldelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardon, D.; Ermolaeva, M.D.; Vamthyan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, R.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A/Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A/Reference number: AB2035; MUID:2040683; PMID:10952301

A/Accession: H82435

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-241 <HEI>

A/Cross-references: UNIPROT:Q9KLV7; UNIPARC:UPI00000C35C1; GB:AE004393; GB:AE003853; NID

A/Experimental source: serogroup O1, strain N16961, biotype El Tor

C/Genetics:

A/Genes: VCA0634

A/Map position: 2

Query Match 40.7%; Score 44; DB 2; Length 241;

Best Local Similarity 52.6%; Pred. No. 17;

Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 ROQIALGAVFESALVDK 21

DB 135 ROQIALGAVFESALVDK 153

RESULT 10

T44255
probable amino acid oxidase flavoprotein [imported] - *Rhizobium etli* plasmid b

C/Species: *Rhizobium etli*

C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C/Accession: T44255

R/Miranda-Rios, J.; Moreira, C.; Taboada, H.; Davalos, A.; Encarnacion, S.; Mora, J.; So

J. Bacteriol. 179, 6887-6893, 1997

A/Title: Expression of chitin biosynthetic genes (*chiCGS*) and production of symbiotic

A/Reference number: Z22737; MUID:98037482; PMID:9371431

A/Accession: T44255

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-327 <MIR>

A/Cross-references: UNIPROT:Q34292; UNIPARC:UPI0000136E97; EMBL:AF004408; NID:G2627325;

A/Experimental source: strain CE3

C/Genetics:

A/Genes: *thiO*

C/Superfamily: D-amino-acid oxidase

Query Match 40.7%; Score 44; DB 2; Length 327;

Best Local Similarity 55.6%; Pred. No. 23;

Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 DPROQIALGAVFESALV 18

DB 138 DPROQIALGAVFESALV 155

RESULT 11

AB3438
amidophosphoribosyltransferase (EC 2.4.2.14) [imported] - *Brucella melitensis* (strain 1

C/Species: *Brucella melitensis*

C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C/Accession: AB3438

R/DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanova,

A/Title: The genome sequence of the facultative intracellular pathogen *Brucella melit*

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Reference number: AD3252; PMID:1175688

A/Accession: AB3438

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-516 <KUR>

A/Cross-references: UNIPROT:Q8YFN2; UNIPARC:UPI00000580B6; GB:AE008917; PIDD:AAL52669.1

A/Experimental source: strain 16M

C/Genetics:

A/Map position: 1

C/Superfamily: amidophosphoribosyltransferase

C/Keywords: glycosyltransferase; pentosyltransferase

Query Match 40.7%; Score 44; DB 2; Length 516;

Best Local Similarity 53.3%; Pred. No. 38;

Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 ROQIALGAVFESALV 17

DB 153 ROQIALGAVFESALV 167

RESULT 12

UT0742

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 Seconds
(without alignments)
267,334 Million cell updates/sec

Title: US-10-501-838A-6
Perfect score: 108
Sequence: 1 DPROQIALGAVFESALVDKLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	206	1 Y2627_PSEAR	Q91011 pseudomonas
2	79.5	73.6	207	2 Q41ZC6_AZOVI	Q41ZC6 azotobacter
3	69	63.9	206	1 Y3359_PSESM	Q47125 pseudomonas
4	69	63.9	206	2 Q4ZRXJ_PSEBSY	Q88568 pseudomonas
5	69	63.9	208	1 Y4015_PSEBS	Q43901 pseudomonas
6	59	54.6	207	2 Q4K9U1_PSEBS	Q51814 pseudomonas
7	50	46.3	582	2 Q5LS14_SILPO	Q91468 nephroblastoma
8	49	45.4	372	2 Q9T468_NEPOL	Q91468 nephroblastoma
9	49	45.4	548	2 Q74DU9_GEOSL	Q74DU9 geobacter
10	48	44.4	606	2 Q22137_ARATH	Q22137 arabidopsis
11	48	44.4	1074	2 Q4FXU8_LEIMA	Q4FXU8 leishmania
12	47.5	44.0	903	2 Q6ANH7_DESPS	Q6ANH7 desulfohalobium
13	47	43.5	381	2 Q54331_STRLI	Q54331 streptomyces
14	47	43.5	381	2 Q9RLY3_STRCO	Q9RLY3 streptomyces
15	47	43.5	594	2 Q92TV7_RHIME	Q92TV7 rhizobium
16	47	43.5	888	2 Q5A114_CANAL	Q5A114 candida alb
17	47	43.5	888	2 Q9HEP0_CANAL	Q9HEP0 candida alb
18	47	43.5	1007	2 Q76H00_MAI2S	Q76H00 zea mays
19	47	43.5	1494	2 Q9XMA6_CAEEL	Q9XMA6 caenorhabditis
20	46.5	43.1	892	2 Q67RX2_SYMT	Q67RX2 bartonella
21	46	42.6	107	2 Q93F32_BAHEH	Q93F32 bartonella
22	46	42.6	126	2 Q84VL1_ARALY	Q84VL1 arabidopsis
23	46	42.6	126	2 Q84UX3_ARATH	Q84UX3 arabidopsis
24	46	42.6	281	2 Q7Y189_ORISA	Q7Y189 oryza sativa
25	46	42.6	326	2 Q7NVJ2_GLOVI	Q7NVJ2 gloeobacter
26	46	42.6	345	2 Q9AT00_ARATH	Q9AT00 arabidopsis
27	46	42.6	364	2 Q80812_ARATH	Q80812 arabidopsis
28	46	42.6	407	2 Q4NW88_ENTPT	Q4NW88 enterococcus
29	46	42.6	492	2 Q834M1_ENTPT	Q834M1 enterococcus
30	46	42.6	546	2 Q4K4T6_MACFA	Q4K4T6 macaca fasc
31	46	42.6	643	1 QPGH_XANCV	Q83442 xanthomonas

RESULT 1	ID	Y2627_PSEAR	STANDARD;	PRT;	206 AA.
AC	Q91011				
DT	10-OCT-2003	(Rel. 42, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DT	10-MAY-2005	(Rel. 47, Last annotation update)			
DE	Hypothetical UPF0274 protein PA2627.				
GN	OrderedLocustNames=PA2627;				
OS	Pseudomonas aeruginosa.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;				
OC	Pseudomonadaceae; Pseudomonas.				
OX	NCBI_TaxID=287;				
RN	[1]				
RP	NCBIOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].				
RC	STRAIN=ATCC 15692 / PA01;				
RX	MDL:1000000000:20437337; PubMed=10984043; DOI=10.1038/35023079;				
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mitsuoguchi S.D., Warren P.,				
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,				
RA	Garber R.L., Goltz L., Tolentino B., Westbrock-Wadman S., Yuan Y.,				
RA	Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,				
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,				
RA	Reizer J., Sailer M.H. Jr., Hancock R.B.W., Lory S., Olson M.V.;				
RT	*Complete genome sequence of Pseudomonas aeruginosa PA01, an				
RT	opportunistic pathogen.				
RT	Nature 406:895-904 (2000)				
CC	SwissProt entry belongs to the UPF0274 family.				
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use as long as its content is in no way modified and this statement is not				
CC	removed.				
CC	EMBL: AB004691; AAC06015.1; -; Genomic DNA.				
DR	PIR: D83317; D83317.				
DR	HANAB: MF_00695; -; 1.				
DR	InterPro: IPR007451; DUF489.				
DR	Pfam: PF04356; DUF489; 1.				
KW	Complete proteome; Hypothetical protein.				
SO	SEQUENCE 206 AA; 23063 MW; F79FPL2B6803F16 CRC64;				

Query Match 100.0%; Score 108; DB 1; Length 206;
Best local similarity 100.0%; Pred. No. 3e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 DPROQIALGAVFESALVDKLA 23
|||||
Db 3 DPROQIALGAVFESALVDKLA 25

RESULT 2
Q41ZC6_AZOVI
ID Q41ZC6_AZOVI PRELIMINARY; PRT; 207 AA.

```

AC 041ZC6;
DT 13-SEP-2005 (TRENBLREL. 31, Created)
DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLREL. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AVINDRAF14589;
OS Azotobacter vinelandii AVOP.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=322710;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Giavina T.,
RA Hammon N., Istrati S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Azotobacter vinelandii
RT AVOP."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Azotobacter vinelandii
RT AVOP."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute;
RA Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Giavina T.,
RA Hammon N., Istrati S., Pittluck S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC DR EMBL; AAU0300003; BAM06292.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 207 AA; 23629 MW; 0928B1PFB908A855 CRC64;

Query Match 73.6%; Score 79.5; DB 2; Length 207;
Best Local Similarity 78.3%; Pred. No. 0.00015;
Matches 18; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 2 PRO-QUILGAVPESALVNDKLA 23
DB 3 PROBOLIALGALFESALVNDRIA 25

RESULT 3
Y3359_PSESM STANDARD; PRT; 206 AA.
AC Q872K5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein PSPT03359.
GN OrderedLocusNames=PSPT03359;
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DC3000;
RG MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,

```

```

RA Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA Nelson M.C., Davidson T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA Utecherback T.R., Van Aken S.E., Feldblum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramoa A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RT Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186 (2003).
CC -!- SIMILARITY: Belongs to the UPF0274 family.
CC CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC DR EMBL; AE016853; AA056837.1; -; Genomic_DNA.
CC DR TIGR; PSPT03359; -;
CC DR HAMAP; MF_00695; -; 1.
CC DR InterPro; IPR007451; DUF489.
CC DR Pfam; PF04356; DUF489; 1.
CC KW Complete proteome; Hypothetical protein.
CC SEQUENCE 206 AA; 22954 MW; BFEDEB85CB742F33 CRC64;

Query Match 63.9%; Score 69; DB 1; Length 206;
Best Local Similarity 61.9%; Pred. No. 0.0084;
Matches 13; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 3 RQOLIALGAVPESALVNDKLA 23
DB 5 RQOLIALGAVPESALVNDRIA 25

RESULT 4
Q4ZRJ9_PSESY PRELIMINARY; PRT; 206 AA.
AC Q4ZRJ9;
DT 13-SEP-2005 (TRENBLREL. 31, Created)
DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLREL. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Pepr_3191;
OS Pseudomonas syringae pv. syringae B728a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=205918;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RG DOE Joint Genome Institute;
RA Chain P., Larimer F., DiBartolo G., Copeland A., Iykidis A., Trong S.,
RA Nolan M., Goltzman E., Thiel J., Malfatti S., Lapidus A., Dettler J.C.,
RA Land M., Richardson P.M., Kyriades N.C., Ivanova N.;
RT "Comparison of two complete genome sequences of Pseudomonas syringae
RT pv. syringae B728a and pv. tomato DC3000."
RT Proc. Natl. Acad. Sci. U.S.A. 0:0-0 (2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RG Loper J.;
RT Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RG Fell H., Fell W.S., Lindow S.B.;
RA Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC DR EMBL; CP000075; AA138223.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 206 AA; 22912 MW; 2180772F6F08DB42 CRC64;

```

Query Match 63.9%; Score 69; DB 2; Length 206;
 Best Local Similarity 61.9%; Pred. No. 0.0084;
 Matches 13; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 3 RQOLIALGAVPESALVDKLA 23
 DB 5 OBOILALGVFOAVALVDRIA 25

RESULT 5

Y4015_PSEBP STANDARD; PRT; 208 AA.
 AC Q88FR8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein PP4015.
 GN OrderedLocustNames=PP4015;
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 CC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=160488;

RM NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RX MEDLINE=22423060; PubMed=12534463;
 RX DOI=10.1046/j.1462-2920.2002.00366.x;
 RA Nelson K.B., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L.M., Beanan M.J., Deboy R.T., Daugherty S.C., Kolonay J.P.,
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khoult H.M.,
 RA Hance I., Chits Lee P., Holtzaple B.K., Scanlan D., Tran K.,
 RA Moazzar A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
 RA Wedler H., Lauber J., Stjepandic D., Hobeisel J., Straetz M., Heim S.,
 RA Kowitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tsemmler B.,
 RA Fraser C.M.,
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440."
 RL Enzymol. Microbiol. 4:799-808(2002).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC -----
 CC EMBL; AE016788; AAN69609.1; -; Genomic_DNA.
 DR TIGR; PP4015; -;
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 208 AA; 23368 MW; 435653B443D2538 CRC64;

Query Match 63.9%; Score 69; DB 1; Length 208;
 Best Local Similarity 61.9%; Pred. No. 0.0085;
 Matches 13; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 3 RQOLIALGAVPESALVDKLA 23
 DB 5 OBOILALGVFOAVALVDRIA 25

RESULT 6

Q4K9U1_PSEPS PRELIMINARY; PRT; 207 AA.
 ID Q4K9U1;
 AC Q4K9U1;
 DT 13-SEP-2005 (TRENBLrel. 31, Created)
 DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=PPL_3892;
 OS Pseudomonas fluorescens (strain Pf-5).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=220664;

RM NUCLEOTIDE SEQUENCE.

RP STRAIN=PF-5;
 RX PubMed=15980861; DOI=10.1038/nbt1110;
 RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
 RA Mavrodil D., Deboy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
 RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
 RA Gwinn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
 RA Khoult H.M., Pierson B., Pierson L., Thomas L., Loper J.,
 RT "Complete genome sequence of the plant commensal Pseudomonas
 RT fluorescens Pf-5."
 RT Nat. Biotechnol. 23:873-878(2005).
 RL EMBL; CP000076; AAY93156.1; -; Genomic_DNA.
 DR HAMAP; PF000916; Sulfate_transp; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 207 AA; 23131 MW; D9B9C6B0CFC3293 CRC64;

Query Match 54.6%; Score 59; DB 2; Length 207;
 Best Local Similarity 57.1%; Pred. No. 0.38;
 Matches 12; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 RQOLIALGAVPESALVDKLA 23
 DB 5 OBOILALGVFOAVALVDRIA 25

RESULT 7

OSL514_SILPO PRELIMINARY; PRT; 582 AA.
 ID OSL514;
 AC OSL514;
 DT 01-FEB-2005 (TRENBLrel. 29, Created)
 DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
 DE Sulfate permease.
 GN Name=slp; OrderedLocustNames=SP01956;
 OS Silicibacter pomeroyi.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
 CC Rhodobacteraceae; Silicibacter.
 NCBI_TaxID=89184;

RM NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RP STRAIN=DS-3 / ATCC 700808 / DSM 15171;
 RX PubMed=15602564; DOI=10.1038/nature03170;
 RA Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
 RA Brinkac L.M., Lewis M., Johnson J.R., King G.M., Belas R., Fugua C.,
 RA Kane E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
 RA Paulsen I.T., Ren Q., Daugherty S.C., Deboy R.T., Dodson R.J.,
 RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
 RA Haft D.H., Selengut J., Ward N.,
 RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
 RT marine environment."
 RL Nature 432:910-913(2004).
 DR EMBL; CP000031; AAY95232.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008271; F:sulfate porter activity; IEA.
 DR GO; GO:0008272; P:sulfate transport; IEA.
 DR InterPro; IPR002645; STAS.
 DR InterPro; IPR011547; Sulpn_transp.
 DR InterPro; IPR001902; Sulp_transp.
 DR Pfam; PF01740; STAS; 1.
 DR Pfam; PF00916; Sulfate_transp; 1.
 DR TIGRFAMs; TIGR00815; slp; 1.
 DR PROSITE; PS50801; STAS; 1.
 KM Complete proteome.
 SQ SEQUENCE 582 AA; 61574 MW; 949A630C17B89B31 CRC64;

Query Match 46.3%; Score 50; DB 2; Length 582;
 Best Local Similarity 61.1%; Pred. No. 0.33;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

QY      1 DPROQIALGAVESAL.18
      |||:|||||:|:|
Db      307 DPROQIALGAVANISSL 324

RESULT 8
Q9T468_NEPOL PRELIMINARY; PRT; 372 AA.
AC      Q9T468;
DT      01-MAY-2000 (TrEMBLrel. 13, Created)
DT      01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE      01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE      Putative plastid division protein.
GN      Name:efw;
OS      Nephroselmis olivacea.
OC      Chloroplast.
OC      Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
OC      Pseudocourfieldiales; Pycnoccocaceae; Nephroselmis.
OX      NCBI_TaxID=31312;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=9398694; PubMed=10468594; DOI=10.1073/pnas.96.18.10248;
RA      Turnel M., Ols C., Lemieux C.;
RT      "The complete chloroplast DNA sequence of the green alga Nephroselmis
RT      olivacea: insights into the architecture of ancestral chloroplast
RT      genomes.";
RL      Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).
DR      EMBL; AF137379; AAD54904.1; -; Genomic DNA.
DR      EMBL; AF137379; AAD54885.1; -; Genomic DNA.
DR      GO; GO:0009507; Chloroplast; IEA.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0007049; P:cell cycle; IEA.
DR      InterPro; IPR001182; Cell cycle.
DR      Pfam; PF01098; FTSW_R0DA_SPOVE; 1.
DR      PROSITE; PS00428; FTSW_R0DA_SPOVE; 1.
KW      Chloroplast.
SQ      SEQUENCE 372 AA; 41969 MW; 5B6A29D2B0E87D22 CRC64;

Query Match      45.4%; Score 49; DB 2; Length 372;
Best Local Similarity 69.2%; Pred. No. 31;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 DPROQIALGAVF 13
      |||:|||||:|:|
Db      302 DPROQIALGAVSLF 314

RESULT 9
Q74DU9_GEOSL PRELIMINARY; PRT; 548 AA.
AC      Q74DU9;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Transport ATP-binding protein CydC.
GN      Name:cydC; OrderedlocusNames=GSU1216;
OS      Geobacter sulfurreducens.
OC      Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC      Geobacteraceae; Geobacter.
OX      NCBI_TaxID=35554;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RX      STRAIN=PCA / ATCC 51573;
RX      PubMed=14671304; DOI=10.1126/science.1089727;
RA      Meithe B.A., Nelson K.E., Eissen J.A., Paulsen I.T., Nelson W.C.,
RA      Heideberg J.F., Wu D., Ward N.L., Beaman M.J., Dodson R.J.,
RA      Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA      Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA      Davidsen T.M., Zafar N., White O., Tran B., Romero C., Fotherger H.A.,
RA      Weidman J.F., Khoult H.M., Feldblyum T.V., Ustebayrak T.R.,
RA      Van Aken S.B., Lovley D.R., Frazer C.M.;
RT      "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT      environments.";

```

```

RL      Science 302:1967-1969 (2003).
DR      EMBL; AE017180; AAR34592.1; -; Genomic DNA.
DR      TIGR; GSU1216; -.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0016887; F:ATPase activity; IEA.
DR      GO; GO:0046265; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR      GO; GO:0000166; F:nucleotide binding; IEA.
DR      GO; GO:0006810; P:transport; IEA.
DR      InterPro; IPR003593; AAA_ATPase.
DR      InterPro; IPR011527; ABC_membrane_1.
DR      InterPro; IPR003439; ABC_transp_like.
DR      InterPro; IPR002453; Beta_tubulin.
DR      Pfam; PR00005; ABC_tran; 1.
DR      ProDom; PD000006; ABC_transporter; 1.
DR      SMART; SM00382; AAA; 1.
DR      PROSITE; PS50929; ABC_TM1F; 1.
DR      PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR      PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR      PROSITE; PS00228; TUBULIN_B_AUTOREG; UNKNOWN_1.
KW      ATP-binding; Complete proteome.
SQ      SEQUENCE 548 AA; 57118 MW; 257422201BD634A CRC64;

Query Match      45.4%; Score 49; DB 2; Length 548;
Best Local Similarity 55.0%; Pred. No. 46;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      4 DQQLIAGVAVESALVDKLA 23
      ::|||:|||||:|:|
Db      210 BELIILGAVEROASIVDDLS 229

RESULT 10
O22137_ARATH PRELIMINARY; PRT; 606 AA.
AC      O22137;
DT      01-JAN-1998 (TrEMBLrel. 05, Created)
DT      01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein At2g45350.
GN      Name=At2g45350;
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC      rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX      NCBI_TaxID=3702;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,
RA      Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R.,
RA      Venter J.C.;
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RA      Town C.D., Kaul S.;
RL      Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AC002387; AAB82628.1; -; Genomic DNA.
DR      PIR; B84889; B84889.
DR      GO; GO:0006381; P:mRNA editing; IMP.
DR      InterPro; IPR002885; PPR.
DR      InterPro; IPR011990; TPR-1like_helical.
DR      Pfam; PR01535; PPR; 11.
DR      TIGRPFAM; TIGR00756; PPR; 10.
KW      Hypothetical protein; Repeat.
SQ      SEQUENCE 606 AA; 68379 MW; E52AD100A3F0F134 CRC64;

Query Match      43.4%; Score 48; DB 2; Length 606;
Best Local Similarity 43.5%; Pred. No. 74;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      1 DPROQIALGAVESALVDKLA 23
      |||:|||||:|:|
Db      94 DPROQIALIICLMLGVSVDFPS 116

```

RESULT 11
 04PXU8 LEIMA
 ID 04PXU8_LEIMA PRELIMINARY; PRT; 1074 AA.
 AC 04PXU8;
 DT 13-SEP-2005 (TREMBLrel. 31, Created)
 DT 13-SEP-2005 (TREMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=LMJ_0747;
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OC NCBI_TaxId=5664;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=Friedlin.
 RA Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
 RA Barriman M., Slick B., Rajandream M.-A., Aert R., Anupama A.,
 RA Apostolou Z., Attipoe P., Baeson N., Bauser C., Beck A., Beverly S.M.,
 RA Blanchettin G., Borzym K., Boche G., Brusch C.V., Collins M.,
 RA Cdag B., Charlont I., Clayton C., Coulson R.M.R., Cronin A.,
 RA Cruz A.K., Davies R.M., De Gaudenzi U., Dobson D.B., Duesterhoeft A.,
 RA Fazelin G., Foster N., Fraehn A.C., Fraser A., Fuchs M., Gabel C.,
 RA Gobie A., Goffeau A., Harris D., Hertz-Powier C., Hilbert H., Horn D.,
 RA Huang Y., Klages S., Knight A., Kube M., Larke N., Litvin L.,
 RA Lord A., Louie T., Marra M., Masny D., Mathews K., Michalek S.,
 RA Motttram J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,
 RA Oliver K., O'Neill S., Pentony M., Pohl T.M., Price C., Purnelle B.,
 RA Quail M.A., Rabinowitsch E., Reinhardt R., Rieger M., Rinta J.,
 RA Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
 RA Schaefer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
 RA Shin H., Sivam D., Squares R., Squares S., Toaso V., Vogt C.,
 RA Volckert G., Wamburt R., Warren T., Wedler H., Woodward J., Zhou S.,
 RA Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
 RA Myler P.U.;
 RT "The genome of the kinetoplastid parasite, Leishmania major."
 RL Science 0:0-(2005).
 DR EMBL/CP000080/AZ09562.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 1074 AA; 10916 MW; 7265ED0833A292D CRC64;
 Query Match 44.4%; Score 48; DB 2; Length 1074;
 Best Local Similarity 42.9%; Pred. No. 1.3e+02;
 Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 Oy 1 DPROQILN-ALGAVPESALVDK 21
 Db 285 DPAPQITLALREMAAQR 305
 RESULT 12
 06ANH7 DESPS
 ID 06ANH7_DESPS PRELIMINARY; PRT; 903 AA.
 AC 06ANH7;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Probable pyruvate phosphate dikinase.
 GN OrderedLocustNames=DP1369;
 OS Desulfococcus psychrophila.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
 OC Desulfobulbaceae; Desulfococcales.
 OC NCBI_TaxId=84980;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=LSV54 / DSM 12343;
 RA Pubmed=15305914; DOI=10.1111/j.1462-2920.2004.00665.x;
 RA Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
 RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gallner K.,
 RA Teeling H., Leuchner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
 RA Klenk H.-P.;
 RT "The genome of Desulfococcus psychrophila, a sulfate-reducing bacterium

RT from permanently cold Arctic sediments.";
 RL Environ. Microbiol. 6:887-902(2004).
 DR EMBL/CR522870; CAG36097.1; -; Genomic_DNA.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0016301; F:kinase activity; IEA.
 DR GO; GO:0000287; F:magnesium ion binding; IEA.
 DR GO; GO:0050242; F:pyruvate, phosphate dikinase activity; IEA.
 DR GO; GO:0016310; F:phosphorylation; IEA.
 DR InterPro; IPR008279; PEP_mobile.
 DR InterPro; IPR000121; PEP_utilizers.
 DR InterPro; IPR002192; PEP_N term.
 DR InterPro; IPR010121; PyrN_phos_dikin.
 DR Pfam; PF00391; PEP_utilizers_I.
 DR Pfam; PF02896; PEP-utilizers_C; 1.
 DR Pfam; PF01326; PEP_N; 1.
 DR Prodom; PD000940; PEP_utilizers; 1.
 DR TIGRPFAMS; TIGR01828; PyrN_phos_dikin; 1.
 DR PROSITE; PS00742; PEP_ENZYMES_2; 1.
 DR PROSITE; PS00370; PEP_ENZYMES_PHOS_SITE; UNKNOWN_1.
 KW Complete proteome; Kinase; Pyruvate.
 SQ SEQUENCE 903 AA; 10062 MW; 0586872F4D6127F CRC64;
 Query Match 44.0%; Score 47.5; DB 2; Length 903;
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;
 Matches 10; Conservative 5; Mismatches 0; Indels 1; Gaps 1;
 Oy 1 DPROQILN-ALGAVPES 15
 Db 213 DPROQLMGSIGAVPKS 228
 RESULT 13
 054331 STRLI
 ID 054331_STRLI PRELIMINARY; PRT; 381 AA.
 AC 054331;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Cellulase B precursor.
 GN Name=celB;
 OS Streptomyces lividans.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.
 OC NCBI_TaxId=1916;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=66;
 RA Wittmann S., Sharek F., Kluepfel D., Morosoli R.,
 RT "Purification and characterization of the CelB endoglucanase from
 RT Streptomyces lividans 66 and DNA sequence of the encoding gene."
 RL Appl. Environ. Microbiol. 60:1701-1703(1994).
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=66;
 RA Sharek F.;
 RL EMBL/U04629; AAB71950.1; -; Genomic_DNA.
 DR PDB; INLR; X-ray; A=41-274.
 DR PDB; 2NLR; X-ray; A=41-274.
 DR GO; GO:0008810; F:cellulase activity; IEA.
 DR GO; GO:0030247; F:polysaccharide binding; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0000272; P:polysaccharide catabolism; IEA.
 DR InterPro; IPR001919; Bac_cellose-bind.
 DR InterPro; IPR01291; Carb_Cell bind.
 DR InterPro; IPR002594; Glyco_hydro_12.
 DR Pfam; PF00553; GMW_2; 1.
 DR Pfam; PF01670; Glyco_hydro_12; 1.
 DR Prodom; PD004316; Glyco_hydro_12; 1.
 DR SMART; SMO0637; CBD II; 1.
 DR PROSITE; PS00561; CBD_BACTERIAL; UNKNOWN_1.
 KW Signal.

```
FT SIGNAL 1 40 Potential.
FT CHAIN 41 381 cellulase-B.
SQ SEQUENCE 381 AA; 39239 MW; A7B99BF590A24EC CRC64;

Query Match 43.5%; Score 47; DB 2; Length 381;
Best Local Similarity 57.1%; Pred. No. 68;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 PROQQLALGAVFESALVDKL 22
DB 11 PRGILAAAGVLAAPALVSSL 31

RESULT 14
Q9RYJ3_STRCO
ID Q9RYJ3_STRCO PRELIMINARY; PRT; 381 AA.
AC Q9RYJ3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative secreted cellulase B.
GN OrderedLocNames=SC01187; ORFNames=SC011A.18;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxId=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RA MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.B., Quail M.A., Kleeser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,
RA Huang C.-H., Kleeser T., Larke L., Murphy U.D., Oliver K., O'Neill S.,
RA Rabinowitz B., Rajadream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wierozorek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939108; CAB61599.1; -; Genomic_DNA.
DR HSSP; OS4331; 2NLR.
DR SMR; O9RYJ3; 41-262.
DR GO; GO:0008810; F:cellulase activity; IEA.
DR GO; GO:0030247; F:polysaccharide binding; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0005972; P:polysaccharide catabolism; IEA.
DR InterPro; IPR001919; Bac cellulose-bind.
DR InterPro; IPR012291; Carb cell bind.
DR InterPro; IPR002594; Glyco_hydro_12.
DR Pfam; PF00553; CBM_2; 1.
DR Pfam; PF01670; Glyco_hydro_12; 1.
DR ProDom; PD004316; Glyco_hydro_12; 1.
DR SMART; SM00637; CBD II; 1.
DR PROSITE; PS00561; CBD_BACTERIAL; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 381 AA; 39199 MW; 97CD8FF58679B4EC CRC64;

Query Match 43.5%; Score 47; DB 2; Length 381;
Best Local Similarity 57.1%; Pred. No. 68;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 PROQQLALGAVFESALVDKL 22
DB 11 PRGILAAAGVLAAPALVSSL 31

RESULT 15
Q92TV7_RHIME
ID Q92TV7_RHIME PRELIMINARY; PRT; 594 AA.
AC Q92TV7;
```

```
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative glyoxylate carboxyligase protein (EC 4.1.1.47).
GN Name=gcl; OrderedLocNames=Rb1388; ORFNames=SM020681;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymB.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxId=382;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;
RA Finn T.M., Meldner S., Wong K., Buhrmester J., Chai P.,
RA Vorhoefer F.O., Hernandez-Lucas I., Becker A., Gouzy J.,
RA Golding B., Puhler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL591985; CAC49788.1; -; Genomic_DNA.
DR PIR; D96015; D96015.
DR HSSP; P07342; INOH.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0009028; P:tartarate-semialdehyde synthase activity; IEA.
DR GO; GO:0030976; P:thiamin pyrophosphate binding; IEA.
DR GO; GO:0009436; P:glyoxylate catabolism; IEA.
DR InterPro; IPR006397; Glyox_carbo_1lg.
DR InterPro; IPR000399; TPP binding.
DR InterPro; IPR011766; TPP enzyme C.
DR InterPro; IPR012000; TPP enzyme M.
DR InterPro; IPR012001; TPP enzyme N.
DR PANTHER; PTHR18968:SF60; Glyox_carbo_1lg; 1.
DR Pfam; PF02775; TPP_enzyme_C; 1.
DR Pfam; PF0205; TPP_enzyme_M; 1.
DR Pfam; PF02776; TPP_enzyme_N; 1.
DR PIRSF; PIRSF00107; G1c1; 1.
DR PIRSF; PIRSF001370; Thdp depend ac1; 1.
DR TIGRFAMs; TIGR01504; glyox_carbo_1lg; 1.
KW Complete proteome; Ligase; Lyase; Plasmid.
SQ SEQUENCE 594 AA; 64583 MW; D6B5ABD1A080EA1F CRC64;

Query Match 43.5%; Score 47; DB 2; Length 594;
Best Local Similarity 34.8%; Pred. No. 11e+02;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 DPROQLALGAVFESALVDKLA 23
DB 435 DPNRQVALSGDYDFQFLIEELA 457
```

Search completed: January 23, 2006, 09:31:11
Job time : 62.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 Seconds
(without alignments)
114.551 Million cell updates/sec

Title: US-10-501-838a-6
Sequence: 1 DPROQIALGAVFESALVDKLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5 COMB.pep.*
2: /cgn2_6/prodata/1/1aa/6 COMB.pep.*
3: /cgn2_6/prodata/1/1aa/H COMB.pep.*
4: /cgn2_6/prodata/1/1aa/RTUS COMB.pep.*
5: /cgn2_6/prodata/1/1aa/RE COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfillseq.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	495	2	US-09-252-991A-31450
2	47	43.5	381	2	US-09-216-295-22
3	47	43.5	381	2	US-09-632-570-22
4	47	43.5	381	2	US-09-632-575-52
5	47	43.5	465	2	US-09-248-796A-17155
6	47	43.5	3129	2	US-09-482-788-2
7	46	42.6	441	2	US-09-328-352-5065
8	46	42.6	493	2	US-09-134-000C-5547
9	44	40.7	535	2	US-09-902-540-15138
10	43	39.8	102	2	US-09-048-502-1
11	43	39.8	124	1	US-08-434-705B-16
12	43	39.8	124	1	US-09-086-201-16
13	43	39.8	129	1	US-08-434-705B-4
14	43	39.8	129	1	US-09-086-201-4
15	43	39.8	383	2	US-09-489-039A-9926
16	43	39.8	622	2	US-09-252-991A-27587
17	42	38.9	146	2	US-09-640-211A-1032
18	42	38.9	166	2	US-09-228-986-113
19	42	38.9	166	2	US-10-101-464A-113
20	42	38.9	304	2	US-09-695-458-3
21	42	38.9	331	2	US-09-695-458-2
22	42	38.9	399	2	US-09-252-991A-20610
23	42	38.9	481	2	US-09-252-991A-27323
24	42	38.9	790	2	US-09-543-681A-27609
25	42	38.9	817	2	US-09-252-991A-27609
26	42	38.9	896	2	US-09-538-092-142
27	41.5	38.4	878	2	US-08-941-936-2

28	41.5	38.4	953	1	US-08-500-857A-2	Sequence 2, Appl1
29	41.5	38.4	955	1	US-08-500-857A-10	Sequence 10, Appl1
30	41.5	38.4	1048	2	US-09-921-099A-11	Sequence 11, Appl1
31	38.0	38.0	69	2	US-09-489-847-156	Sequence 156, Appl1
32	38.0	38.0	86	2	US-09-621-976-4833	Sequence 4833, Appl1
33	38.0	38.0	198	2	US-10-104-047-3702	Sequence 3702, Appl1
34	38.0	38.0	205	2	US-09-489-039A-8206	Sequence 8206, Appl1
35	38.0	38.0	232	2	US-09-902-540-11664	Sequence 11664, Appl1
36	38.0	38.0	307	2	US-09-252-991A-24189	Sequence 24189, Appl1
37	38.0	38.0	386	2	US-09-248-796A-16103	Sequence 16103, Appl1
38	38.0	38.0	427	2	US-09-252-991A-31877	Sequence 31877, Appl1
39	38.0	38.0	460	2	US-09-252-991A-23461	Sequence 23461, Appl1
40	38.0	38.0	489	2	US-09-991-181-138	Sequence 138, Appl1
41	38.0	38.0	489	2	US-09-990-444-138	Sequence 138, Appl1
42	38.0	38.0	489	2	US-09-997-333-138	Sequence 138, Appl1
43	38.0	38.0	489	2	US-09-992-598-138	Sequence 138, Appl1
44	38.0	38.0	505	2	US-09-729-995-2	Sequence 2, Appl1
45	38.0	38.0	505	2	US-09-729-995-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-252-991A-31450
Sequence 31450, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubinfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31450
LENGTH: 495
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31450

Query Match 100.0%; Score 108; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 2.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPROQIALGAVFESALVDKLA 23
DB 292 DPROQIALGAVFESALVDKLA 314

RESULT 2
US-09-216-295-22
Sequence 22, Application US/09216295
Patent No. 6268328
GENERAL INFORMATION:
APPLICANT: Mitchinson, Colin
TITLE OF INVENTION: No. 6268328el Variant EGIII-like Cellulase Compositions
FILE REFERENCE: GC555
CURRENT APPLICATION NUMBER: US/09/216,295
CURRENT FILING DATE: 1998-12-18
NUMBER OF SEQ ID NOS: 41
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 22
LENGTH: 381
TYPE: PRT
ORGANISM: Streptomyces lividans Celb
US-09-216-295-22

```
Query Match          43.5%; Score 47; DB 2; Length 381;
Best Local Similarity 57.1%; Pred. No. 7.8;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Oy 2 PROQIILAGVAFESALVDKL 22
   |||||:|||||
   11 PRGLIILAGVILAFALVSSL 31
Db

RESULT 3
US-09-632-570-22
; Sequence 22, Application US/09632570
; Patent No. 6623949
; GENERAL INFORMATION:
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Phillips, Jay Ian
; TITLE OF INVENTION: No. 6623949el Variant EGIII-like Cellulase
; FILE REFERENCE: GCG31
; CURRENT APPLICATION NUMBER: US/09/632,570
; CURRENT FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Streptomyces lividans CelB
US-09-632-570-22

Query Match          43.5%; Score 47; DB 2; Length 381;
Best Local Similarity 57.1%; Pred. No. 7.8;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Oy 2 PROQIILAGVAFESALVDKL 22
   |||||:|||||
   11 PRGLIILAGVILAFALVSSL 31
Db

RESULT 4
US-09-632-575-52
; Sequence 52, Application US/09632575
; Patent No. 6635465
; GENERAL INFORMATION:
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Ropp, Tracy M.
; TITLE OF INVENTION: Mutant EGIII Cellulase, DNA Encoding
; FILE REFERENCE: GCG29
; CURRENT APPLICATION NUMBER: US/09/632,575
; CURRENT FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Streptomyces lividans CelB
US-09-632-575-52

Query Match          43.5%; Score 47; DB 2; Length 381;
Best Local Similarity 57.1%; Pred. No. 7.8;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Oy 2 PROQIILAGVAFESALVDKL 22
   |||||:|||||
   11 PRGLIILAGVILAFALVSSL 31
Db

RESULT 5
US-09-248-796A-17155
; Sequence 17155, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:
; APPLICANT: Ketch Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17155
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17155

Query Match          43.5%; Score 47; DB 2; Length 465;
Best Local Similarity 46.2%; Pred. No. 9.9;
Matches 12; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

Oy 1 DPROQIILAG-----AVFESALVD 20
   :|||:|||||
   394 EPRQVLETLGLIAGLVISFESAGTVD 419
Db

RESULT 6
US-09-482-788-2
; Sequence 2, Application US/09482788
; Patent No. 6893839
; GENERAL INFORMATION:
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; APPLICANT: Wendy T. Yoder
; TITLE OF INVENTION: Methods For Producing Polypeptides In
; FILE REFERENCE: 5778.200-US
; CURRENT APPLICATION NUMBER: US/09/482,788
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: 09/229,862
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3129
; TYPE: PRT
; ORGANISM: Fusarium
US-09-482-788-2

Query Match          43.5%; Score 47; DB 2; Length 3129;
Best Local Similarity 43.5%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Oy 1 DPROQIILAGVAFESALVDKLA 23
   |||||:|||||
   1897 DPROQVIGIGVMBELVVTGDLA 1919
Db

RESULT 7
US-09-328-352-5065
; Sequence 5065, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5065
; LENGTH: 441
```


TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-5065

Query Match 42.6%; Score 46; DB 2; Length 441;
Best Local Similarity 36.4%; Pred. No. 14;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 DPROQIALGAVFESALVDKL 22
| | | | | : | | : | | : |
DB 386 DARGELVLGMDQALQRL 407

RESULT 8
US-09-134-000C-5547
Sequence 5547, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5547
LENGTH: 493
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-5547

Query Match 42.6%; Score 46; DB 2; Length 493;
Best Local Similarity 52.4%; Pred. No. 16;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 RQOLIALGAVFESALVDKLA 23
| | | | | : | | : | | : |
DB 269 KERLIKQAGVFLSPALADKLA 289

RESULT 9
US-09-902-540-15138
Sequence 15138, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 15138
LENGTH: 535
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-15138

Query Match 40.7%; Score 44; DB 2; Length 535;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 QIALGAVFESALVD 20
| | | | | : | | : | | : |
DB 367 QVAVGMAALQSHALD 382

RESULT 10
US-09-048-502-1
Sequence 1, Application US/09048502
Patent No. 6458927
GENERAL INFORMATION:

APPLICANT: Hestrup, Svend
APPLICANT: Christensen, Kenneth
APPLICANT: Thim, Lars
APPLICANT: Judge, Martin Edward
APPLICANT: Kristensen, Peter
TITLE OF INVENTION: Polypeptide With Appetite Regulating
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 64589270 No. 6458927disk of No. 6458927th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/048,502
FILING DATE: 26-MAR-1998
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 4618,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-09-048-502-1

Query Match 39.8%; Score 43; DB 2; Length 102;
Best Local Similarity 52.4%; Pred. No. 7.6;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PROQIALGAVFESALVDKL 22
| | | | | : | | : | | : |
DB 27 PRQLRAGVAVLQIALQEV 47

RESULT 11
US-08-434-705B-16
Sequence 16, Application US/08434705B
Patent No. 5798258
GENERAL INFORMATION:

APPLICANT: Douglas Ph.D., James
TITLE OF INVENTION: CART PROTEIN AND DNA ENCODING THEREFOR
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis K. Shelton
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/434,705B
FILING DATE: May 4, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: OHSU18446
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-8100
TELEFAX: (206) 224-0779
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
US-08-434-705B-16

Query Match 39.8%; Score 43; DB 1; Length 124;
Best Local Similarity 52.4%; Pred. No. 9.7;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 PROQIALGAVFESALVDKL 22
DB 41 PRROLRAPGAVLQIEALQEV 61

RESULT 12
US-09-086-201-16
Sequence 16, Application US/09086201
Patent No. 5969103
GENERAL INFORMATION:
APPLICANT: Douglas Ph.D., James
TITLE OF INVENTION: CART PROTEIN AND DNA ENCODING THEREFOR
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis K. Shelton
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/086,201
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/434,705
FILING DATE: May 4, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: OHSU18446
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-8100
TELEFAX: (206) 224-0779
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 124 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

DESCRIPTION: PET23b-rCART2 fusion protein
HYPOTHETICAL: YES
US-09-086-201-16

Query Match 39.8%; Score 43; DB 1; Length 124;
Best Local Similarity 52.4%; Pred. No. 9.7;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 PROQIALGAVFESALVDKL 22
DB 41 PRROLRAPGAVLQIEALQEV 61

RESULT 13
US-08-434-705B-4
Sequence 4, Application US/08434705B
Patent No. 5798258
GENERAL INFORMATION:
APPLICANT: Douglas Ph.D., James
TITLE OF INVENTION: CART PROTEIN AND DNA ENCODING THEREFOR
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis K. Shelton
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,705B
FILING DATE: May 4, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: OHSU18446
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-8100
TELEFAX: (206) 224-0779
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 129 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-434-705B-4

Query Match 39.8%; Score 43; DB 1; Length 129;
Best Local Similarity 52.4%; Pred. No. 10;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 PROQIALGAVFESALVDKL 22
DB 54 PRROLRAPGAVLQIEALQEV 74

RESULT 14
US-09-086-201-4
Sequence 4, Application US/09086201
Patent No. 5969103
GENERAL INFORMATION:
APPLICANT: Douglas Ph.D., James
TITLE OF INVENTION: CART PROTEIN AND DNA ENCODING THEREFOR
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis K. Shelton
STREET: 1420 Fifth Avenue, Suite 2800
CITY: Seattle

STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/086,201
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/434,705
FILING DATE: May 4, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: OHSU18446
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-8100
TELEFAX: (206) 224-0779
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 129 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-09-086-201-4

Query Match 39.8%; Score 43; DB 1; Length 129;
Best local Similarity 52.4%; Pred. No. 10;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PROQIATGAVFESALVDKL 22
DB 54 PRQLRAPGAVLQIRALQEVLT 74

RESULT 15
US-09-489-039A-9926
Sequence 9926, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 9926
LENGTH: 383
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9926

Query Match 39.8%; Score 43; DB 2; Length 383;
Best local Similarity 42.1%; Pred. No. 39;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 3 RQQLIAGAVFESALVDK 21
DB 122 RRERRAIGTVFQSSALLSR 140

Search completed: January 23, 2006, 09:35:46
Job time : 16.6 secs

This Page Blank (unused)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838a-6
Perfect score: 108
Sequence: 1 DPROQIALGAVESALVDKLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications, AA, Main:*

- 1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	23	US-10-665-184-6	Sequence 6, Appl1
2	108	100.0	23	US-10-942-300-6	Sequence 6, Appl1
3	108	100.0	23	US-10-501-838a-6	Sequence 6, Appl1
4	108	100.0	206	US-10-665-184-64	Sequence 64, Appl1
5	108	100.0	206	US-10-942-300-64	Sequence 64, Appl1
6	108	100.0	206	US-10-501-838a-58	Sequence 58, Appl1
7	49	45.4	208	US-10-437-863-173815	Sequence 173815, Appl1
8	49	45.4	2353	US-10-437-863-137865	Sequence 137865, Appl1
9	48	44.4	606	US-10-740-618-2	Sequence 2, Appl1
10	47	43.5	206	US-10-425-115-365459	Sequence 365459, Appl1
11	47	43.5	271	US-10-767-701-39716	Sequence 39716, Appl1
12	47	43.5	381	US-10-441-625-22	Sequence 22, Appl1
13	47	43.5	381	US-10-441-625-22	Sequence 22, Appl1
14	47	43.5	888	US-10-389-566-1057	Sequence 1057, Appl1
15	46	42.6	345	US-10-696-616-30	Sequence 30, Appl1
16	46	42.6	357	US-10-739-930-9819	Sequence 9819, Appl1
17	46	42.6	632	US-10-418-8618-69	Sequence 69, Appl1
18	45.5	42.1	334	US-10-282-122A-67178	Sequence 67178, Appl1
19	45.5	42.1	334	US-10-406-686A-44	Sequence 44, Appl1
20	45	41.7	173	US-10-450-763-39406	Sequence 39406, Appl1
21	45	41.7	311	US-10-425-115-221896	Sequence 221896, Appl1
22	45	41.7	381	US-10-425-115-221898	Sequence 221898, Appl1
23	44	40.7	307	US-10-156-761-11838	Sequence 11838, Appl1
24	44	40.7	352	US-10-450-763-55845	Sequence 55845, Appl1
25	44	40.7	458	US-10-282-122A-53974	Sequence 53974, Appl1
26	44	40.7	618	US-10-369-993-19204	Sequence 19204, Appl1
27	44	40.7	878	US-10-450-763-55850	Sequence 55850, Appl1

28	43.5	40.3	438	US-11-097-143-35538	Sequence 35538, A
29	43	39.8	102	US-09-048-502-1	Sequence 1, Appl1
30	43	39.8	102	US-10-228-591-1	Sequence 1, Appl1
31	43	39.8	242	US-10-425-115-294570	Sequence 294570, Appl1
32	43	39.8	437	US-10-369-493-2821	Sequence 2821, Appl1
33	43	39.8	591	US-10-437-963-143877	Sequence 143877, Appl1
34	43	39.8	603	US-10-425-114-68167	Sequence 68167, Appl1
35	43	39.8	633	US-10-437-963-143875	Sequence 143875, Appl1
36	43	39.8	724	US-10-732-923-10675	Sequence 10675, A
37	43	39.8	785	US-10-425-115-257218	Sequence 257218, A
38	42.5	39.4	543	US-10-282-122A-74785	Sequence 74785, A
39	42	38.9	77	US-10-424-599-178640	Sequence 178640, A
40	42	38.9	87	US-09-978-360A-472	Sequence 472, Appl1
41	42	38.9	140	US-10-425-115-192048	Sequence 192048, Appl1
42	42	38.9	146	US-10-856-493-1032	Sequence 1032, Appl1
43	42	38.9	154	US-10-264-049-3399	Sequence 3399, Appl1
44	42	38.9	166	US-10-101-464A-113	Sequence 113, Appl1
45	42	38.9	166	US-10-864-252-113	Sequence 113, Appl1

ALIGNMENTS

RESULT 1
US-10-665-184-6
Sequence 6, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 6
LENGTH: 23
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-665-184-6
Query Match 100.0%; Score 108; DB 4; Length 23;
Best Local Similarity 100.0%; Pred No. 6, 4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 DPROQIALGAVESALVDKLA 23
1 DPROQIALGAVESALVDKLA 23
RESULT 2
US-10-942-300-6
Sequence 6, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-503
CURRENT APPLICATION NUMBER: US/10/942,300
PRIOR FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 6
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-10-942-300-6

Query Match 100.0%; Score 108; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 6,4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DPROQLALGAVFESALVDKLA 23
Db 1 DPROQLALGAVFESALVDKLA 23

RESULT 3
US-10-501-838A-6
; Sequence 6, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-501-838A-6

Query Match 100.0%; Score 108; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 6,4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DPROQLALGAVFESALVDKLA 23
Db 1 DPROQLALGAVFESALVDKLA 23

RESULT 4
US-10-665-184-64
; Sequence 64, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 64
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-10-665-184-64

Query Match 100.0%; Score 108; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DPROQLALGAVFESALVDKLA 23
Db 3 DPROQLALGAVFESALVDKLA 25

RESULT 5
US-10-942-300-64

; Sequence 64, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 64
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-942-300-64

Query Match 100.0%; Score 108; DB 5; Length 206;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DPROQLALGAVFESALVDKLA 23
Db 3 DPROQLALGAVFESALVDKLA 25

RESULT 6
US-10-501-838A-58
; Sequence 58, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-501-838A-58

Query Match 100.0%; Score 108; DB 5; Length 206;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPROQIALGAVFESALVDKLA 23
 Db 3 DPROQIALGAVFESALVDKLA 25

RESULT 7
 US-10-437-963-173915
 ; Sequence 173915, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Bardazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 173915
 ; LENGTH: 208
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_71907C.1.pep
 US-10-437-963-173915

Query Match 45.4%; Score 49; DB 4; Length 208;
 Best Local Similarity 36.8%; Pred. No. 7.6; 4; Indels 0; Gaps 0;
 Matches 7; Conservative 8; Mismatches 0;
 Qy 1 DPROQIALGAVFESALV 19
 Db 108 EPHREVALIGRIFFESGKLM 126

RESULT 8
 US-10-437-963-137865
 ; Sequence 137865, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Bardazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 137865
 ; LENGTH: 2353
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_39307C.1.pep
 US-10-437-963-137865

Query Match 45.4%; Score 49; DB 4; Length 2353;
 Best Local Similarity 36.8%; Pred. No. 1.3e+02; 4; Indels 0; Gaps 0;
 Matches 7; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DPROQIALGAVFESALV 19
 Db 2266 EPHREVALIGRIFFESGKLM 2284

RESULT 9
 US-10-740-618-2
 ; Sequence 2, Application US/10740618
 ; Publication No. US20050026280A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TASAKA, MASAO
 ; APPLICANT: SHINKAWA, TOSHIHARU
 ; TITLE OF INVENTION: GENE INVOLVED IN CHOROPLAST RNA-EDITING
 ; FILE REFERENCE: 246948USO
 ; CURRENT APPLICATION NUMBER: US/10/740,618
 ; CURRENT FILING DATE: 2003-12-22
 ; PRIOR APPLICATION NUMBER: JP2003-199098
 ; PRIOR FILING DATE: 2003-07-18
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2
 ; LENGTH: 606
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-10-740-618-2

Query Match 44.4%; Score 48; DB 5; Length 606;
 Best Local Similarity 43.5%; Pred. No. 39;
 Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 DPROQIALGAVFESALVDKLA 23
 Db 94 DPROQIALGAVFESALVDKLA 116

RESULT 10
 US-10-425-115-365459
 ; Sequence 365459, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 365459
 ; LENGTH: 206
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_96464C.1.pep
 US-10-425-115-365459

Query Match 43.5%; Score 47; DB 4; Length 206;
 Best Local Similarity 44.0%; Pred. No. 16;
 Matches 11; Conservative 6; Mismatches 4; Indels 4; Gaps 1;

Qy 2 PROQIALGAVFESALVDKLA 22
 Db 125 PROQIALGAVFESALVDKLA 149

RESULT 11
 US-10-767-701-39716
 ; Sequence 39716, Application US/10767701
 ; Publication No. US20040172684A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 39716
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C28675_1.pep
US-10-767-701-39716

Query Match      43.5%; Score 47; DB 4; Length 271;
Best Local Similarity 45.5%; Pred. No. 23;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Cy      1 DPROQIALGAVFESALVDKL 22
Db      161 DEDMEGRIGLVFQSAALPDSL 182

RESULT 12
US-10-441-626-22
; Sequence 22, Application US/10441626
; Publication No. US20030186418A1
; GENERAL INFORMATION:
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Phillips, Jay Ian
; TITLE OF INVENTION: No. US20030186418A1e1 Variant EGIII-Like Cellulase
; FILE REFERENCE: GC631
; CURRENT APPLICATION NUMBER: US/10/441,626
; CURRENT FILING DATE: 2003-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Streptomyces lividans CelB
US-10-441-626-22

Query Match      43.5%; Score 47; DB 4; Length 381;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Cy      2 PROQIALGAVFESALVDKL 22
Db      11 PRGLAALGAVLAAPALVSSL 31

RESULT 13
US-10-441-625-22
; Sequence 22, Application US/10441625
; Publication No. US20030203467A1
; GENERAL INFORMATION:
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Phillips, Jay Ian
; TITLE OF INVENTION: Novel Variant EGIII-Like Cellulase
; FILE REFERENCE: GC631
; CURRENT APPLICATION NUMBER: US/10/441,625
; CURRENT FILING DATE: 2003-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
```

```

; ORGANISM: Streptomyces lividans CelB
US-10-441-625-22

Query Match      43.5%; Score 47; DB 4; Length 381;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Cy      2 PROQIALGAVFESALVDKL 22
Db      11 PRGLAALGAVLAAPALVSSL 31

RESULT 14
US-10-389-566-1057
; Sequence 1057, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1057
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-389-566-1057

Query Match      43.5%; Score 47; DB 4; Length 888;
Best Local Similarity 46.2%; Pred. No. 89;
Matches 12; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

Cy      1 DPROQIALG-----AVFESALVD 20
Db      817 EPRQVLETLGLAGLVSIFFSAGTVD 842

RESULT 15
US-10-696-616-30
; Sequence 30, Application US/10696616
; Publication No. US20040191850A1
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua
; APPLICANT: Tossberg, John
; APPLICANT: Zhou, Qing
; APPLICANT: McElver, John
; APPLICANT: Aux, George
; APPLICANT: Budziszewski, Greg
; APPLICANT: Thomas, Carla
; APPLICANT: Patton, David
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Essential for Plant Grow
; FILE REFERENCE: 70166USNP
; CURRENT APPLICATION NUMBER: US/10/696,616
; CURRENT FILING DATE: 2003-10-29
; PRIOR APPLICATION NUMBER: US 60/423,519
; PRIOR FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
```


US-10-696-616-30

Query Match 42.6%; Score 46; DB 4; Length 345;

Best Local Similarity 52.9%; Pred. No. 44;

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 7 TATGAVFESALVDKLA 23

Db 160 LRIGLVFQSAALFDSLS 176

Search completed: January 23, 2006, 11:52:33
Job time : 61 secs

This Page Blank (verso)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 Seconds
(without alignments)
114.551 Million cell updates/sec

Title: US-10-501-838a-7

Perfect score: 104
Sequence: 1 LIDNRVLAAGVQALQVROIA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patente AA: *
1: /cgn2_6/ptodata/1/1aa/5.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/6.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/H.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/PCYTUS.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/RE.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	48.1	265	2 US-09-328-352-4347	Sequence 4347, Ap
2	50	48.1	678	2 US-09-252-991A-20693	Sequence 20693, A
3	49	47.1	289	2 US-09-328-352-5601	Sequence 5601, Ap
4	48	46.2	216	2 US-09-543-681A-7993	Sequence 7993, Ap
5	48	46.2	217	2 US-09-489-039A-7789	Sequence 7789, Ap
6	47	45.2	147	2 US-09-710-279-850	Sequence 850, App
7	47	45.2	323	2 US-09-134-001C-3339	Sequence 3339, Ap
8	46	44.2	369	2 US-09-487-558B-268	Sequence 268, App
9	45	43.3	143	2 US-09-513-999C-7973	Sequence 7973, Ap
10	45	43.3	623	2 US-10-104-047-2547	Sequence 2547, Ap
11	44	42.3	79	2 US-09-107-532A-7073	Sequence 7073, Ap
12	44	42.3	377	2 US-09-487-558B-270	Sequence 270, App
13	44	42.3	522	2 US-09-252-991A-31244	Sequence 31244, A
14	43	41.3	759	2 US-09-902-540-16161	Sequence 16161, A
15	42.5	40.9	589	2 US-09-252-991A-18093	Sequence 18093, A
16	42	40.4	271	2 US-09-248-796A-18309	Sequence 18309, A
17	42	40.4	456	1 US-08-910-731-6	Sequence 6, Appl1
18	42	40.4	461	1 US-08-910-731-6	Sequence 2816, Ap
19	42	40.4	466	2 US-09-605-703B-2816	Sequence 19728, A
20	42	40.4	511	2 US-09-252-991A-19728	Sequence 30529, A
21	42	40.4	602	2 US-09-252-991A-30529	Sequence 102, App
22	42	40.4	841	2 US-09-533-029-102	Sequence 27923, A
23	41.5	39.9	397	2 US-09-252-991A-27923	Sequence 701, App
24	41	39.4	205	2 US-09-640-211A-701	Sequence 26340, A
25	41	39.4	291	2 US-09-252-991A-26340	Sequence 1337, Ap
26	41	39.4	297	2 US-09-538-032-1337	Sequence 24201, A
27	41	39.4	725	2 US-09-252-991A-24201	

28	41	39.4	832	2 US-09-252-991A-19252	Sequence 19252, A
29	41	39.4	1021	2 US-09-489-039A-13589	Sequence 13589, A
30	41	39.4	2871	2 US-09-538-092-936	Sequence 936, App
31	40	38.5	50	2 US-10-729-121-32	Sequence 32, Appl
32	40	38.5	121	2 US-09-134-000C-4714	Sequence 4714, Ap
33	40	38.5	205	2 US-09-553-662-26	Sequence 26, Appl
34	40	38.5	205	2 US-10-062-994-26	Sequence 26, Appl
35	40	38.5	253	2 US-09-107-433-2831	Sequence 2831, Ap
36	40	38.5	270	2 US-09-252-991A-16844	Sequence 16844, A
37	40	38.5	292	2 US-09-489-039A-1431	Sequence 12431, A
38	40	38.5	299	2 US-09-583-110-4728	Sequence 4728, Ap
39	40	38.5	351	2 US-09-431-976-2	Sequence 2, Appl1
40	40	38.5	351	2 US-09-021-286-2	Sequence 2, Appl1
41	40	38.5	351	2 US-10-729-121-2	Sequence 2, Appl1
42	40	38.5	449	2 US-09-252-991A-27941	Sequence 27941, A
43	40	38.5	465	2 US-09-902-540-12612	Sequence 12612, A
44	40	38.5	521	2 US-09-252-991A-25430	Sequence 25430, A
45	40	38.5	549	2 US-09-252-991A-20196	Sequence 20196, A

ALIGNMENTS

RESULT 1
US-09-328-352-4347
Sequence 4347, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4347
LENGTH: 265
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-4347

Query Match 48.1% Score 50; DB 2; Length 265;
Best Local Similarity 60.0%; Pred. No. 2.4;
Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 NRVLALAGVQALQVROIA 23

DB 38 NRVLALAGVQALQVROIA 57

RESULT 2
US-09-252-991A-20693
Sequence 20693, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20693

LENGTH: 678
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20693

Query Match 48.1% Score 50; DB 2; Length 678;

Query Match 45.2%; Score 47; DB 2; Length 323;
Best Local Similarity 35.0%; Pred. No. 9.5;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LIDNRVIALAGVVOALQOVR 20
Db 220 ILSNKAMALIGVIELLKQAK 239

RESULT 8
US-09-487-5588-268

Sequence 268, Application US/094875588
Patent No. 6949356
GENERAL INFORMATION:
APPLICANT: Bubby, Robert
APPLICANT: Call, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. 6949356man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.130
CURRENT APPLICATION NUMBER: US/09/487,5588
CURRENT FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/487,558
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 446
SOFTWARE: Patentin version 3.0
SEQ ID NO 268
LENGTH: 369
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-487-5588-268

Query Match 44.2%; Score 46; DB 2; Length 369;
Best Local Similarity 30.8%; Pred. No. 16;
Matches 8; Conservative 9; Mismatches 5; Indels 4; Gaps 1;

Qy 1 LIDNRVIALAG---VVOALQOVRQI 22
Db 218 LVNKNIFCLHGLSPMIFETIDVREL 243

RESULT 9
US-09-513-999C-7973

Sequence 7973, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclercq, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961
FILE REFERENCE: 59, US2, REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 7973
LENGTH: 143
TYPE: PRT
ORGANISM: Homo sapiens
US-09-513-999C-7973

Query Match 43.3%; Score 45; DB 2; Length 143;
Best Local Similarity 42.9%; Pred. No. 7.8;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LIDNRVIALAGVVOALQOVRQ 21
Db 48 LVSEKKSALASVKEIKELRQLRQ 68

RESULT 10
US-10-104-047-2547

Sequence 2547, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241el full length cDNA
FILE REFERENCE: HI-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2547
LENGTH: 623
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-2547

Query Match 43.3%; Score 45; DB 2; Length 623;
Best Local Similarity 42.9%; Pred. No. 44;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LIDNRVIALAGVVOALQOVRQ 21
Db 465 LVSEKKSALASVKEIKELRQLRQ 485

RESULT 11
US-09-107-532A-7073

Sequence 7073, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Steamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:

```

;
; TELEPHONE: (781)893-5007
; TELERAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 7073:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 79 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHEICAL: YES
;   ORIGINAL SOURCE:
;     ORGANISM: Enterococcus faecium
;   FEATURE:
;     NAME/KEY: msc feature
;     LOCATION: (B) LOCATION 1...79
;   SEQUENCE DESCRIPTION: SEQ ID NO: 7073:
US-09-107-532A-7073

Query Match          42.3%; Score 44; DB 2; Length 79;
Best Local Similarity 36.4%; Pred. No. 5.7;
Matches 8; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LIDNRVLAAGVQALQOVRQI 22
Db 7 LIDRQAVRIAQVYKAVQKTRAV 28

RESULT 12
US-09-487-558B-270
; Sequence 270, Application US/09487558B
; Patent No. 6949356
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. 6949356man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.130
; CURRENT APPLICATION NUMBER: US/09/487,558B
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/487,558
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 446
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 270
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-487-558B-270

Query Match          42.3%; Score 44; DB 2; Length 377;
Best Local Similarity 30.8%; Pred. No. 35;
Matches 8; Conservative 8; Mismatches 6; Indels 4; Gaps 1;

Qy 1 LIDNRVLAAGVQALQOVRQI 22
Db 226 LVDNKIFCLHGLSPMIEITIDQVRDL 251

RESULT 13
US-09-252-991A-31244
; Sequence 31244, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31244
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31244

Query Match          42.3%; Score 44; DB 2; Length 522;
Best Local Similarity 56.2%; Pred. No. 52;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 7 LALAGVQALQOVRQI 22
Db 92 LAGAAVEALRQVRQV 107

RESULT 14
US-09-902-540-16161
; Sequence 16161, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16161
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16161

Query Match          41.3%; Score 43; DB 2; Length 759;
Best Local Similarity 43.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 LIDNRVLAAGVQALQOVRQI 23
Db 294 LKDRLEWFAFAGIVPALSRVEQDA 316

RESULT 15
US-09-252-991A-18093
; Sequence 18093, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18093
```

```

1  LENGTH: 589
2  TYPE: PRT
3  ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18093

```

	Query Match	40.9%	Score 42.5;	DB 2;	length 589;
	Best Local Similarity	46.2%	Pred. No. 1.le+02;		
	Matches	12;	Conservative	5;	Mismatches 4;
					Indels 5;
					Gaps 1;
Cy	1	LIDRV-----IALAGVQAALQVRG	21		
		::: :	::: :		
Dd	343	LVDHQDVAAVALAVLGVCAAMLVTRG	368		

Search completed: January 23, 2006, 09:35:47
Job time : 17.6 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:56 ; Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838a-7

Perfect score: 104

Sequence: 1 LIDNRVLAAGVVALQGVRLA 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/PCIT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US03_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	47	45.2	147 6 US-10-793-626-850	Sequence 850, App
2	47	45.2	297 6 US-10-878-556A-119	Sequence 119, App
3	42	40.4	743 6 US-10-467-657-4082	Sequence 4082, App
4	40	38.5	50 6 US-10-729-121-32	Sequence 32, App1
5	40	38.5	351 6 US-10-729-121-2	Sequence 2, App1
6	38	36.5	288 6 US-10-873-528-148	Sequence 148, App
7	38	36.5	478 6 US-11-037-829A-4	Sequence 4, App1
8	38	36.5	487 6 US-11-193-473-2	Sequence 2, App1
9	38	36.5	619 6 US-10-821-234-1150	Sequence 1150, App
10	37	35.6	78 6 US-10-467-657-9176	Sequence 9176, App
11	37	35.6	109 6 US-10-821-234-1223	Sequence 1223, App
12	37	35.6	151 6 US-11-103-957-72	Sequence 72, App1
13	37	35.6	162 6 US-10-467-657-1176	Sequence 1176, App
14	37	35.6	244 6 US-10-793-626-566	Sequence 566, App
15	37	35.6	254 6 US-11-067-323-636	Sequence 636, App
16	37	35.6	281 6 US-10-878-556A-77	Sequence 27, App1
17	37	35.6	566 6 US-10-467-657-3302	Sequence 3302, App
18	37	35.6	584 6 US-11-024-959-380	Sequence 380, App
19	37	35.6	753 6 US-10-858-730-225	Sequence 225, App
20	37	35.6	770 6 US-10-982-545-15	Sequence 15, App1
21	37	35.6	770 6 US-10-789-273-38	Sequence 38, App1
22	37	35.6	970 6 US-10-770-726-86	Sequence 86, App1
23	37	35.6	3375 6 US-11-044-111-23	Sequence 23, App1
24	36	34.6	211 6 US-10-793-626-744	Sequence 744, App
25	36	34.6	274 7 US-11-082-389-150	Sequence 150, App

26	36	34.6	289 6 US-10-793-626-64	Sequence 64, App1
27	36	34.6	290 6 US-10-793-626-3024	Sequence 3024, App
28	36	34.6	290 6 US-10-793-626-3252	Sequence 3252, App
29	36	34.6	366 6 US-10-454-437-256	Sequence 256, App
30	36	34.6	366 6 US-11-055-822-364	Sequence 364, App
31	36	34.6	366 6 US-10-979-821-12	Sequence 12, App1
32	36	34.6	391 6 US-11-114-922-12	Sequence 12, App1
33	36	34.6	468 6 US-10-467-657-2274	Sequence 2274, App
34	36	34.6	549 6 US-11-082-489-148	Sequence 148, App
35	36	34.6	751 6 US-11-019-711-115	Sequence 115, App
36	36	34.6	932 7 US-11-057-058-59	Sequence 59, App1
37	36	34.6	1290 6 US-10-485-517-141	Sequence 141, App
38	36	34.6	1604 7 US-11-037-243-73	Sequence 73, App1
39	35	33.7	189 7 US-11-205-109-33	Sequence 184, App
40	35	33.7	209 6 US-10-454-437-184	Sequence 7940, App
41	35	33.7	311 6 US-10-467-657-7940	Sequence 184, App
42	35	33.7	344 6 US-10-454-437-148	Sequence 148, App
43	35	33.7	374 6 US-10-793-626-1664	Sequence 1664, App
44	35	33.7	442 6 US-10-513-639-25	Sequence 25, App1
45	35	33.7	494 7 US-11-024-959-445	Sequence 445, App

ALIGNMENTS

```
RESULT 1
US-10-793-626-850
; Sequence 850, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OR INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCES: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 850
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-850

Query Match 45.2%; Score 47; DB 6; Length 147;
Best local similarity 35.0%; Pred. No. 0.46;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 LIDNRVLAAGVVALQGVRLA 20
   ::::|::|::|::|::|::|
Db 44 ILSNKAMLLGVIELKQAK 63

RESULT 2
US-10-878-556A-119
; Sequence 119, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OR INVENTION: HCV regulated protein expression
; FILE REFERENCES: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 119
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
```

PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: gw_hum/nadc_human
DATABASE ENTRY DATE: 1997-11-01
US-10-878-556A-119

Query Match 45.2%; Score 47; DB 6; Length 297;
Best Local Similarity 47.6%; Pred. No. 1.1;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 DNRVLAGVVALQOVRQ 23
DB 172 DNRVLAGVGVKAVRAAQA 192

RESULT 3
US-10-467-657-4082
Sequence 4082, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SpA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIRZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 4082
LENGTH: 743
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4082

Query Match 40.4%; Score 42; DB 6; Length 743;
Best Local Similarity 47.4%; Pred. No. 25;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 IDNRVLAGVVALQOVR 20
DB 171 VQNRSLSGKVMALRRR 189

RESULT 4
US-10-729-121-32
Sequence 32, Application US/10729121
Publication No. US20040144397A1
GENERAL INFORMATION:
APPLICANT: Conkling, Mark
TITLE OF INVENTION: MODIFYING NICOTINE AND NITROSAMINE
FILE REFERENCE: VTOB_033C1
CURRENT APPLICATION NUMBER: US/10/729,121
CURRENT FILING DATE: 2003-12-04
PRIOR APPLICATION NUMBER: 60/297,154
PRIOR FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: PCTUS02/18040
PRIOR FILING DATE: 2002-06-06
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 32
LENGTH: 50
TYPE: PRT
ORGANISM: Nicotiana tabacum
US-10-729-121-32

Query Match 38.5%; Score 40; DB 6; Length 50;
Best Local Similarity 47.4%; Pred. No. 1.7;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 DNRVLAGVVALQOVRQ 21
DB 23 DNRHISAGGVKALKSVQ 41

RESULT 5
US-10-729-121-2
Sequence 2, Application US/10729121
Publication No. US20040144397A1
GENERAL INFORMATION:
APPLICANT: Conkling, Mark
TITLE OF INVENTION: MODIFYING NICOTINE AND NITROSAMINE
FILE REFERENCE: VTOB_033C1
CURRENT APPLICATION NUMBER: US/10/729,121
CURRENT FILING DATE: 2003-12-04
PRIOR APPLICATION NUMBER: 60/297,154
PRIOR FILING DATE: 2001-06-08
PRIOR APPLICATION NUMBER: PCTUS02/18040
PRIOR FILING DATE: 2002-06-06
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 351
TYPE: PRT
ORGANISM: Nicotiana tabacum
US-10-729-121-2

Query Match 38.5%; Score 40; DB 6; Length 351;
Best Local Similarity 47.4%; Pred. No. 21;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 DNRVLAGVVALQOVRQ 21
DB 208 DNRHISAGGVKALKSVQ 226

RESULT 6
US-10-873-528-148
Sequence 148, Application US/10873528
Publication No. US20050276814A1
GENERAL INFORMATION:
APPLICANT: Microbial Technics Limited
APPLICANT: Gilbert, Christophe FG
APPLICANT: Hanbro, Philip M
TITLE OF INVENTION: Proteins
FILE REFERENCE: PWC/P21129WO
CURRENT APPLICATION NUMBER: US/10/873,528
CURRENT FILING DATE: 2004-06-23
PRIOR APPLICATION NUMBER: US/09/769,787
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: GB 9816337.1
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: US 60/125164
PRIOR FILING DATE: 1999-03-19
NUMBER OF SEQ ID NOS: 388
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 148
LENGTH: 288
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-10-873-528-148

Query Match 36.5%; Score 38; DB 6; Length 288;
Best Local Similarity 44.4%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 LIDNRVLAGVVALQO 18
DB 180 VLSNGVIALAGALIAQOE 197

```
RESULT 7
US-11-037-829A-4
; Sequence 4, Application US/11037829A
; Publication No. US2005025551A1
; GENERAL INFORMATION:
; APPLICANT: Targacept, Inc
; APPLICANT: Catholic Healthcare West
; APPLICANT: Bencherif, Merouane
; APPLICANT: Lukas, Ronald J.
; TITLE OF INVENTION: Methods and Compositions Relating to Chimeric
; TITLE OF INVENTION: Metcortic Receptor Subunits
; FILE REFERENCE: 1103 1520.PCT
; CURRENT APPLICATION NUMBER: US/11/037,829A
; PRIOR FILING DATE: 2005-01-18
; PRIOR APPLICATION NUMBER: US 60/397,380
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapien
US-11-037-829A-4

Query Match          36.5%; Score 38; DB 7; Length 478;
Best Local Similarity 46.7%; Pred. No. 66;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy      7 LALAGVQALQOVR 21
Db      413 LAVGCLQLQLSIRQ 427

RESULT 8
US-11-193-473-2
; Sequence 2, Application US/11193473
; Publication No. US20060003375A1
; GENERAL INFORMATION:
; APPLICANT: MAO, Yunlin
; APPLICANT: XIE, Yi
; TITLE OF INVENTION: A NOVEL POLYPEPTIDE - HUMAN RETINOIC ACID-RESPONSIVE PROTEIN 53.5
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING THE SAME
; FILE REFERENCE: 011241.52054US
; CURRENT APPLICATION NUMBER: US/11/193,473
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/362,240
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: PCT/CN01/01252
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 00119715.0
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent version 3.1
; SEQ ID NO 2
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-193-473-2

Query Match          36.5%; Score 38; DB 7; Length 487;
Best Local Similarity 53.3%; Pred. No. 67;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      6 VIALAGVQALQOVR 20
Db      312 LLLLVGVVPTIQKVR 326

RESULT 9
US-10-821-234-1150
; Sequence 1150, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crafin, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc seq_genes Version 1.0
; SEQ ID NO 1150
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1150

Query Match          36.5%; Score 38; DB 6; Length 619;
Best Local Similarity 50.0%; Pred. No. 92;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy      7 LALAGVQALQOVR 20
Db      34 LAIGGVLRARQIK 47

RESULT 10
US-10-467-657-9176
; Sequence 9176, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 9176
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-9176

Query Match          35.6%; Score 37; DB 6; Length 78;
Best Local Similarity 35.3%; Pred. No. 9.5;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy      7 LALAGVQALQOVR 23
Db      38 LSDGILQCVQDIKHIS 54

RESULT 11
US-10-821-234-1223
; Sequence 1223, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crafin, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
```

```
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ PRIOR FILING DATE: 2003-04-07
/ NUMBER OF SEQ ID NOS: 1704
/ SOFTWARE: pc_seq_genes Version 1.0
/ SEQ ID NO 1223
/ LENGTH: 109
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-821-234-1223
```

```
Query Match      35.6%; Score 37; DB 6; Length 109;
Best Local Similarity 52.9%; Pred. No. 15;
Matches 9; Conservative 4; Mismatches 2; Indels 2; Gaps 1;
```

```
QY      2 IDNRVLAGVVALQOQ 18
          ::||: ||||| |||
Db      26 VENQ-SLRGVDELQO 40
```

```
RESULT 12
US-11-103-957-72
/ Sequence 72, Application US/11103957
```

```
/ Publication No. US20050281847A1
/ GENERAL INFORMATION:
/ APPLICANT: Berthet, Francois-Xavier Jacques
/ APPLICANT: Lobet, Yves
/ APPLICANT: Poolman, Jan
/ APPLICANT: Verlant, Vincent Georges Christian Louis
/ TITLE OF INVENTION: Vaccine Composition
/ FILE REFERENCE: B45261
/ CURRENT APPLICATION NUMBER: US/11/103,957
/ CURRENT FILING DATE: 2005-04-12
/ PRIOR APPLICATION NUMBER: US/10/467,534
/ PRIOR FILING DATE: 2004-02-03
/ PRIOR APPLICATION NUMBER: PCT/EP02/01356
/ PRIOR FILING DATE: 2002-02-08
/ PRIOR APPLICATION NUMBER: GB 0103169.9
/ PRIOR FILING DATE: 2001-02-08
/ NUMBER OF SEQ ID NOS: 108
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 72
/ LENGTH: 151
/ TYPE: PRT
/ ORGANISM: Chlamydia trachomatis
US-11-103-957-72
```

```
Query Match      35.6%; Score 37; DB 7; Length 151;
Best Local Similarity 43.8%; Pred. No. 22;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      3 DNRVLAGVVALQOQ 18
          ||: ||||| |||
Db      127 DERLNAIGVVELSR 142
```

```
RESULT 13
US-10-467-657-1176
```

```
/ Sequence 1176, Application US/10467657
/ Publication No. US20050260581A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SpA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MASTIGNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/467,657
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ PRIOR FILING DATE: 2001-02-12
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqWIn9, version 1.04
```

```
/ SEQ ID NO 1176
/ LENGTH: 162
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1176
```

```
Query Match      35.6%; Score 37; DB 6; Length 162;
Best Local Similarity 42.1%; Pred. No. 24;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      3 DNRVLAGVVALQOVRQ 21
          ||: ||||| |||
Db      81 DFQVLAQAQPIPIESVRQ 99
```

```
RESULT 14
US-10-793-626-566
/ Sequence 566, Application US/10793626
```

```
/ Publication No. US2005025478A1
/ GENERAL INFORMATION:
/ APPLICANT: KIMMERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PU3480US
/ CURRENT APPLICATION NUMBER: US/10/793,626
/ CURRENT FILING DATE: 2004-03-04
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 566
/ LENGTH: 244
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-566
```

```
Query Match      35.6%; Score 37; DB 6; Length 244;
Best Local Similarity 42.9%; Pred. No. 41;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      2 IDNRVLAGVVALQOVRQI 22
          ||: ||||| |||
Db      134 INNRALALAEKMTKXQVDNV 154
```

```
RESULT 15
US-11-067-323-636
/ Sequence 636, Application US/11067323
```

```
/ Publication No. US20050272064A1
/ GENERAL INFORMATION:
/ APPLICANT: Davis, Christopher S.
/ APPLICANT: Fox, Richard J
/ APPLICANT: Gavrilovic, Vesna
/ APPLICANT: Huitema, Gjalte W
/ APPLICANT: Newman, Lisa M
/ TITLE OF INVENTION: Improved Halohydrin Dehalogenases and Related Polynucleotides
/ FILE REFERENCE: 16028US04 0353.410US
/ CURRENT APPLICATION NUMBER: US/11/067,323
/ CURRENT FILING DATE: 2005-02-23
/ PRIOR APPLICATION NUMBER: 60/546,033
/ PRIOR FILING DATE: 2004-02-18
/ PRIOR APPLICATION NUMBER: 60/494,382
/ PRIOR FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: 10/917,179
/ PRIOR FILING DATE: 2004-08-11
/ NUMBER OF SEQ ID NOS: 1143
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 636
/ LENGTH: 254
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
```

FEATURE:
OTHER INFORMATION: HHDI variant
US-11-067-323-636

Query Match 35.6%; Score 37; DB 7; Length 254;
Best Local Similarity 36.1%; Pred. No. 43;
Matches 13; Conservative 4; Mismatches 5; Indels 14; Gaps 2;

QY 1 LIDNRVLAL-----AGVQALQOVROIA 23
|:|:|
|:|:|
Db 76 LVSNIDIALBFPIDKVAVEDYRGVREAL-QVRPPA 110

Search completed: January 23, 2006, 11:53:36
Job time : 5.1 secs

This Page Blank (usps)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-7

Sequence: 1 LIDNRVLAAGVVALQGVQVQIA 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_21:*

1: Geneseqp19808:.*
2: Geneseqp19908:.*
3: Geneseqp20008:.*
4: Geneseqp20018:.*
5: Geneseqp20028:.*
6: Geneseqp20038:.*
7: Geneseqp20048:.*
8: Geneseqp20058:.*
9: Geneseqp20068:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	23	6	ADBI6892 Xylella f
2	104	100.0	23	9	AEBO8224 Xylella f
3	104	100.0	204	9	AEBO8267 Xylella f
4	50	48.1	127	6	ABU40321 Protein e
5	50	48.1	265	6	ADK33060 Acinetoba
6	50	48.1	614	8	AD157757 Human Bre
7	50	48.1	678	7	ABO71947 Pseudomon
8	49	47.1	23	6	ADBI6889 Vibrio ch
9	49	47.1	23	9	AEBO8221 Vibrio ch
10	49	47.1	204	9	AEBO8264
11	49	47.1	289	6	ADBI6886 Haemophil
12	48	46.2	23	6	ADBI6886 Haemophil
13	48	46.2	23	6	ADBI6886 Haemophil
14	48	46.2	23	9	AEBO8218 Haemophil
15	48	46.2	23	9	AEBO8218 Haemophil
16	48	46.2	29	6	ADBI6886 Haemophil
17	48	46.2	29	6	ADBI6886 Haemophil
18	48	46.2	205	4	AAW50230 Penetrati
19	48	46.2	205	9	AEBO8261 Haemophil
20	48	46.2	216	7	ADFO7708 Bacterial
21	48	46.2	217	7	ABO61272 Klebsiell
22	48	46.2	279	3	AAW54754 Arabidops
23	48	46.2	304	3	AAW30063 Arabidops
24	48	46.2	318	3	AAW30062 Arabidops

25	48	46.2	318	8	ADN72465
26	48	46.2	396	6	ABU25798
27	48	46.2	442	4	ABG96469
28	47	45.2	147	4	AAW81878
29	47	45.2	232	8	ABO84908 Human can
30	47	45.2	297	8	ABM81749 Tumour-as
31	47	45.2	297	9	ADM08782 Human pro
32	47	45.2	308	6	ABU43116 Protein e
33	47	45.2	323	5	ABP38494 Staphyloc
34	47	45.2	323	8	ADSO4663 Staphyloc
35	46	44.2	23	6	ADBI6887 Pasteurel
36	46	44.2	203	9	AEBO8219 Pasteurel
37	46	44.2	276	3	AAW06453 Arabidops
38	46	44.2	286	3	AAW06452 Arabidops
39	46	44.2	303	3	AAW52869 Arabidops
40	46	44.2	369	6	ABR53649 Protein s
41	46	44.2	369	7	ADK61998 Disease t
42	46	44.2	369	8	ADN18915 Bacterial
43	46	44.2	143	3	AAW03892 Human sec
44	45	43.3	254	7	ADG76999 Human nuc
45	45	43.3	254	7	ADG76999 Human nuc

ALIGNMENTS

RESULT 1
ADBI6892
ID ADBI6892 standard; peptide; 23 AA.
XX
AC ADBI6892;
XX
DT 20-NOV-2003 (first entry)
XX
DE Xylella fastidiosa XF1439 penetrating peptide 7.
XX
KW Penetrating peptide; epithelial; endothelial; tight junction; diabetes;
KW Infertility; hormone; vitamin deficiency; neurodegenerative;
KW Cardiovascular; haematological; endocrine disorder; obesity;
KW Neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
KW Osteopathic; cyostatic; nootropic.
XX
OS Xylella fastidiosa.
XX
PN WO2003066859-A2.
XX
PD 14-AUG-2003.
XX
PP 07-FEB-2003; 2003WO-IB000968.
XX
PR 07-FEB-2002; 2002US-0355396P.
XX
PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Ben-Saason SA, Cohen E.
XX
DR WPI; 2003-697452/66.
XX
PT New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
PS Claim 2; Page 14; 60pp; English.
XX
CC This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
CC cytotactic or neurotropic activities. This peptide is from XF1439 of
CC Xylella fastidiosa and is penetrating peptide 7 of the invention.
SQ Sequence 23 AA;
Query Match 100.0%; Score 104; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LIDNRVLTALAGVQALQVQRQIA 23
1 LIDNRVLTALAGVQALQVQRQIA 23
DB 1 LIDNRVLTALAGVQALQVQRQIA 23
RESULT 2
AEB08224 standard; peptide; 23 AA.
AC AEB08224;
DT 25-AUG-2005 (first entry)
XX Xylella fastidiosa XF1439 penetrating peptide 7, SEQ ID NO: 7.
XX
XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
XX antidiabetic; endocrine disease; gastrointestinal disease;
XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
XX neurotropic; neurological disease; parkinsons disease; antiparkinsonian;
XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
XX autoimmune disease; immunosuppressive; immune deficiency;
XX anticonvulsant; genetic disorder; cardiovascular disease;
XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
XX genitourinary disease; hematological disease; antihaemic; anemia;
XX autoimmune disease; immunosuppressive; immune deficiency;
XX immunostimulant; infectious disease; antimicrobial; infection;
XX erectile dysfunction; andrology; major depressive disorder;
XX antidepressant; psychiatric disorder; pain; analgesic;
XX bacterial infection; antibacterial; viral infection; virucide;
XX fungal infection; fungicide; parasitic infection; antiparasitic;
XX renal failure; antifertility; antirheumatic; cytostatic;
XX antiinflammatory; hepatotropic; hepatitis B virus infection.
XX Xylella fastidiosa.
XX
XX US2005136103-A1.
XX
XX 23-JUN-2005.
XX
XX 16-SEP-2004; 2004US-00942300.
XX
XX 17-SEP-2003; 2003US-00664989.
XX
XX 17-SEP-2003; 2003US-00665184.
XX
XX 17-SEP-2003; 2003US-0503615P.
XX
XX (BENS/) BEN-SASSON S A.
XX
XX (COHE/) COHEN E.
XX
XX Ben-Sasson SA, Cohen E;
XX
XX WPI, 2005-444089/45.
XX
XX Composition used for translocating effectors across barrier such as
XX epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
XX
XX Claim 46; SEQ ID NO 7; 59pp; English.
XX
XX The present invention relates to a pharmaceutical composition of
XX penetrating peptides for transcellular delivery of effector. The
XX invention comprises the effector sequentially coupled with a counter ion
XX and at least one hydrophobic agent, where the effector is selectively
XX encapsulated into a complex. The invention is useful for translocating
XX effectors across a biological barrier such as epithelial cells and
XX endothelial cells during treatment and prevention of disease or
XX pathological conditions (including endocrine disorders, diabetes,
XX infertility, hormone deficiencies, osteoporosis, ophthalmological
XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
XX Parkinson's disease, multiple sclerosis, Huntington's disease,
XX cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
XX coagulable states, coronary disease, cerebrovascular events, metabolic
XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
XX hematological disorders, anemia of different entities, immunologic and
XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
XX infectious diseases, viral infections, bacterial infections, fungal
XX infections, parasitic infections, neoplastic diseases, multi-factorial
XX disorders, impotence, chronic pain, depression, different fibrosis states
XX and short stature) and for mucosal vaccination against anthrax and
XX hepatitis B. The present sequence is the Xylella fastidiosa XF1439
XX penetrating peptide. This sequence is used in the effective translocation
XX of antimicrobial antibiotics and antifungal agents across an epithelial
XX barrier.
SQ Sequence 23 AA;
Query Match 100.0%; Score 104; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LIDNRVLTALAGVQALQVQRQIA 23
1 LIDNRVLTALAGVQALQVQRQIA 23
DB 1 LIDNRVLTALAGVQALQVQRQIA 23
RESULT 3
AEB08267 standard; protein; 204 AA.
XX AEB08267;
AC AEB08267;
DT 25-AUG-2005 (first entry)
XX
XX Xylella fastidiosa protein, SEQ ID NO: 65.
XX
XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
XX antidiabetic; endocrine disease; gastrointestinal disease;
XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
XX neurotropic; neurological disease; parkinsons disease; antiparkinsonian;
XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
XX autoimmune disease; immunosuppressive; immune deficiency;
XX anticonvulsant; genetic disorder; cardiovascular disease;
XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
XX genitourinary disease; hematological disease; antihaemic; anemia;
XX autoimmune disease; immunosuppressive; immune deficiency;
XX immunostimulant; infectious disease; antimicrobial; infection;
XX erectile dysfunction; andrology; major depressive disorder;
XX antidepressant; psychiatric disorder; pain; analgesic;
XX bacterial infection; antibacterial; viral infection; virucide;
XX fungal infection; fungicide; parasitic infection; antiparasitic;
XX renal failure; antifertility; antirheumatic; cytostatic;
XX antiinflammatory; hepatotropic; hepatitis B virus infection.
XX Xylella fastidiosa.
XX
XX

PN US2005136103-A1.
XX
XX 23-JUN-2005.
XX
PF 16-SEP-2004; 2004US-00942300.
XX
XX 17-SEP-2003; 2003US-00664989.
XX 17-SEP-2003; 2003US-00665184.
PR 17-SEP-2003; 2003US-0503615P.
XX
XX (BENS/) BEN-SASSON S A.
PA (COHE/) COHEN E.
XX
XX Ben-Sasson SA, Cohen E;
PI
XX WPI; 2005-444089/45.
XX
XX Composition used for translocating effectors across barrier such as
PT epithelial cells during treatment of e.g. endocrine disorders comprises
PT effector sequentially coupled with counter ion and hydrophobic agent.
XX
XX Disclosure; SEQ ID NO 65; 59pp; English.
PS
XX The present invention relates to a pharmaceutical composition of
CC penetrating peptides for transepithelial delivery of effector. The
CC invention comprises the effector sequentially coupled with a counter ion
CC and at least one hydrophobic agent, where the effector is selectively
CC encapsulated into a complex. The invention is useful for translocating
CC effectors across a biological barrier such as epithelial cells and
CC endothelial cells during treatment and prevention of disease or
CC pathological conditions (including endocrine disorders, diabetes,
CC infertility, hormone deficiencies, osteoporosis, ophthalmological
CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
CC Parkinson's disease, multiple sclerosis, Huntington's disease,
CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
CC coagulable states, coronary disease, cerebrovascular events, metabolic
CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC hematological disorders, anemia of different entities, immunologic and
CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
CC infectious diseases, viral infections, bacterial infections, multi-factorial
CC infections, parasitic infections, neoplastic diseases, multi-factorial
CC disorders, impotence, chronic pain, depression, different fibrosis states
CC and short stature) and for mucosal vaccination against anthrax and
CC hepatitis B. The present sequence is the Xylella fastidiosa protein
CC containing penetrating peptide at N-terminal end.
XX
XX Sequence 204 AA;
SQ
Query Match 100.0%; Score 104; DB 9; Length 204;
Best Local Similarity 100.0%; Pred. No. 2,7e-08;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LIDNRVTLALAGVVALQGVROIA 23
DB 4 LIDNRVTLALAGVVALQGVROIA 26
RESULT 4
ABU40321
ID ABU40321 standard; protein; 127 AA.
XX
XX AC
XX ABU40321;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #25848.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Pseudomonas putida.
XX
XX WO200277183-A2.
XX

PD 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX
XX N-PSDB; ACA44191.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 68245; 176pp; English.
PS
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 127 AA;
SQ
Query Match 48.1%; Score 50; DB 6; Length 127;
Best Local Similarity 61.1%; Pred. No. 6.7;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 IDNRVTLALAGVVALQGV 19
DB 83 IDNRVTLALAGVVALQGV 100
RESULT 5
ADA33060
ID ADA33060 standard; protein; 265 AA.
XX
XX AC
XX ADA33060;
XX

DT 20-NOV-2003 (first entry)
 XX Acinetobacter baumannii protein #221.
 XX
 XX Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KM plant biocontrol agent.
 XX
 XX Acinetobacter baumannii.
 XX
 XX US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 XX
 PP 04-JUN-1999; 99US-00328352.
 XX
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX Breton G, Bush D;
 PI
 XX MPI, 2003-576092/54.
 DR N-PSDB; ADA28934.
 XX
 XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 PS Example; SEQ ID NO 4347; 328bp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 CC
 SQ Sequence 265 AA;
 XX
 XX
 Query Match 48.1%; Score 50; DB 6; Length 265;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 4 NRVTLAAGVVALQOVROIA 23
 DB 38 NRVTLAAGVVALQOVROIA 57
 XX
 RESULT 6
 ADI57757
 ID ADI57757 standard; protein; 614 AA.
 XX
 XX ADI57757;
 AC
 XX 22-APR-2004 (first entry)
 DT
 XX Human breast specific protein (BSP) #34.
 DE
 XX Human; breast specific protein; BSP; metastasis; breast cancer;
 KM cytotoxic.
 XX
 XX Homo sapiens.
 OS
 XX MO2003106648-A2.
 PN
 XX 24-DEC-2003.
 PD
 XX 16-JUN-2003; 2003MO-US018934.
 PF
 XX

PR 14-JUN-2002; 2002US-0389327P.
 XX
 XX (DIAD-) DINDEXUS INC.
 PA
 XX Salceda S, Macina RA, Turner LR, Sun Y, Liu C;
 PI
 XX MPI; 2004-082185/08.
 DR N-PSDB; ADI57685.
 XX
 XX Novel isolated polypeptide comprising breast specific protein sequences,
 PT useful for diagnosing or monitoring presence and metastases of breast
 PT cancer in patient.
 PT
 PS Claim 12; SEQ ID NO 128; 370pp; English.
 XX
 XX The invention relates to human breast specific nucleic acids (BSNA) and
 CC the breast specific proteins (BSP) they encode. The nucleic acids are
 CC useful for determining the presence of a BSNA in a sample which involves
 CC contacting the sample with a BSNA under conditions in which the BSNA will
 CC selectively hybridize to a BSNA in the sample, and detecting the
 CC hybridization. The nucleic acids are useful for determining the presence
 CC of a BSP in a sample which involves contacting the sample with suitable
 CC reagent under conditions in which the reagent will selectively interact
 CC with the BSP, and detecting the interaction of the reagent with a BSP in
 CC the sample. The nucleic acids and proteins are useful for diagnosing or
 CC monitoring the presence and metastases of breast cancer in a patient.
 CC which involves determining an amount of nucleic acid or protein in the sample
 CC comprising the determined amount of nucleic acid or protein in the sample
 CC of the patient to the amount of a breast specific marker in a normal
 CC control, where a difference in the determined amount in the sample
 CC compared to the amount in the control is associated with the presence of
 CC breast cancer. The sequences are useful for treating a patient with
 CC breast cancer, involving administering a composition consisting of a BSNA
 CC or a BSP to a patient, where the administration induces an immune
 CC response against the breast cancer cell expressing the BSNA or BSP. This
 CC sequence represents a human BSP of the invention.
 CC
 SQ Sequence 614 AA;
 XX
 XX
 Query Match 48.1%; Score 50; DB 8; Length 614;
 Best Local Similarity 66.7%; Pred. No. 41;
 Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 2 IDNRVTLAAGVVALQOV 19
 DB 408 IDNRVTLAAGVVALASV 425
 XX
 RESULT 7
 ABO71947
 ID ABO71947 standard; protein; 678 AA.
 XX
 XX ABO71947;
 AC
 XX 29-JUL-2004 (first entry)
 DT
 XX Pseudomonas aeruginosa polypeptide #4122.
 DE
 XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 KM
 XX Pseudomonas aeruginosa.
 OS
 XX US6551795-B1.
 PN
 XX 22-APR-2003.
 PD
 XX 18-FEB-1999; 99US-0025291.
 PF
 XX 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX

PR	17-SEP-2003; 2003US-00664989.
PR	17-SEP-2003; 2003US-00665184.
PR	17-SEP-2003; 2003US-050361SP.
PX	
PA	(BENS/) BEN-SASSON S A.
PA	(COHE/) COHEN E.
P1	Ben-Sasson SA, Cohen E;
DR	WPI; 2005-444089/45.
PT	
XX	
PT	Composition used for translocating effectors across barrier such as
PT	epithelial cells during treatment of e.g. endocrine disorders comprises
PT	effector sequentially coupled with counter ion and hydrophobic agent.
PS	
PS	Claim 46; SEQ ID NO 4; 59pp; English.
XX	
CC	The present invention relates to a pharmaceutical composition of
CC	penetrating peptides for transepithelial delivery of effector. The
CC	invention comprises the effector sequentially coupled with a counter ion
CC	and at least one hydrophobic agent, where the effector is selectively
CC	encapsulated into a complex. The invention is useful for translocating
CC	effectors across a biological barrier such as epithelial cells and
CC	endothelial cells during treatment and prevention of disease or
CC	pathological conditions (including endocrine disorders, diabetes,
CC	infertility, hormone deficiencies, osteoporosis, ophthalmological
CC	disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
CC	cardiovascular disorders, atherosclerosis, Huntington's disease,
CC	coagulable states, coronary disease, cerebrovascular events, metabolic
CC	disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC	hematological disorders, anemia of different entities, immunologic and
CC	hematologic diseases, autoimmune diseases, immune deficiencies,
CC	infectious diseases, viral infections, bacterial infections, fungal
CC	infections, parasitic infections, neoplastic diseases, multi-factorial
CC	disorders, impotence, chronic pain, depression, different fibrosis states
CC	and short stature) and for mucosal vaccination against anthrax and
CC	hepatitis B. The present sequence is the Vibrio cholerae VCII27
CC	penetrating peptide. This sequence is used in the effective translocation
CC	of aminoglycoside antibiotics and antifungal agents across an epithelial
CC	barrier.
XX	
SQ	Sequence 23 AA;
Query Match	47.1%; Score 49; DB 9; Length 23;
Best Local Similarity	45.5%; Pred. No. 1,3;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0	
Oy	2 IDNRVLTAGVVQLQOVROIA 23 : : : : : :
Db	2 IYDRTTAFAGICQAVLVQOVA 23
RESULT 10	
AEB08264	AEB08264 standard; protein; 204 AA.
XX	
AC	AEB08264;
XX	
DT	25-AUG-2005 (first entry)
XX	
DE	Vibrio cholerae protein, SEQ ID NO: 62.
XX	
KM	Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
KM	antidiabetic; endocrine disease; gastrointestinal disease;
KM	metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
KM	degenerative; musculoskeletal disease; ocular disease; ophthalmological;
KM	neurodegenerative disease; neuroprotective; Alzheimers disease;
KM	neurologic; neurological disease; parkinsons disease; antiparkinsonian;
KM	dementia; multiple sclerosis; immune disorder; huntingtons chorea;
KM	anticonvulsant; genetic disorder; cardiovascular disease;
KM	cardiovascular-gen.; atherosclerosis; antiarrhythmic;
KM	coronary artery disease; cardiac; vasotonic; obesity; anorectic;

XX	nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
XX	gastrointestinal disease; hematological disease; anti-anemic; anemia;
XX	autoimmune disease; immunosuppressive; immune deficiency;
XX	immunostimulant; infectious disease; antimicrobial; infection;
XX	excitile dysfunction; andrology; major depressive disorder;
XX	antidepressant; psychiatric disorder; pain; analgesic;
XX	bacterial infection; antibacterial; viral infection; virucide;
XX	fungal infection; fungicide; parasitic infection; antiparasitic;
XX	renal failure; antifertility; antirheumatic; cytostatic;
XX	antiflammatory; hepatotropic; hepatitis B virus infection.
OS	Vibrio cholerae.
XX	
XX	US2005136103-A1.
XX	
PD	23-JUN-2005.
XX	
XX	16-SEP-2004; 2004US-00942300.
PR	17-SEP-2003; 2003US-00664989.
PR	17-SEP-2003; 2003US-00665184.
PR	17-SEP-2003; 2003US-0503615P.
XX	
PA	(BENS/) BEN-SASSON S A.
PA	(COHE/) COHEN E.
P1	Ben-Sasson SA, Cohen E;
DR	WPI; 2005-444089/45.
XX	
PT	Composition used for translocating effectors across barrier such as
PT	epithelial cells during treatment of e.g. endocrine disorders comprises
PT	effector sequentially coupled with counter ion and hydrophobic agent.
XX	
PS	Disclosure; SEQ ID NO 62; 59pp; English.
XX	
XX	The present invention relates to a pharmaceutical composition of
CC	penetrating peptides for transepithelial delivery of effector. The
CC	invention comprises the effector sequentially coupled with a counter ion
CC	and at least one hydrophobic agent, where the effector is selectively
CC	encapsulated into a complex. The invention is useful for translocating
CC	effectors across a biological barrier such as epithelial cells and
CC	endothelial cells during treatment and prevention of disease or
CC	pathological conditions (including endocrine disorders, diabetes,
CC	infertility, hormone deficiencies, osteoporosis, ophthalmological
CC	disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
CC	Parkinson's disease, multiple sclerosis, Huntington's disease,
CC	cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
CC	coagulable states, coronary disease, cerebrovascular events, metabolic
CC	disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC	hematological disorder, anemia of different entities, immunologic and
CC	rheumatologic disorder, autoimmune diseases, immune deficiencies,
CC	infectious diseases, viral infections, bacterial infections, fungal
CC	infections, parasitic infections, neoplastic diseases, multi-factorial
CC	disorders, impotence, chronic pain, depression, different fibrosis states
CC	and short stature) and for mucosal vaccination against anthrax and
CC	hepatitis B. The present sequence is the Vibrio cholerae protein
CC	containing penetrating peptide at N-terminal end.
XX	
XX	Sequence 204 AA:
XX	
Query Match	47.1%; Score 49; DB 9; Length 204;
Best Local Similarity	45.5%; Pred. No. 17;
Matches	10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
Qy	2 IDNFTVLAAGVQALQGVQVQIA 23
	: : : : : :
Db	5 IYDRTIAFGICQVAVLQVQA 26
RESULT 11	
ADA34314	
ID	ADA34314 standard; protein; 289 AA

```

XX AC ADA34314;
XX DT 20-NOV-2003 (first entry)
XX DE Acinetobacter baumannii protein #1475.
XX KM Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
XX KM plant biocontrol agent.
XX OS Acinetobacter baumannii.
XX PN US6562958-B1.
XX PD 13-MAY-2003.
XX PF 04-JUN-1999; 99US-00328352.
XX PR 09-JUN-1998; 98US-0088701P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton G, Bush D;
XX PT WPI; 2003-576092/54.
XX DR N-PSDB; ADA30188.
XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX PT for diagnosing a bacterial disease, as components of antibacterial
XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX PT plants.
XX PS Example; SEQ ID NO 5601; 328pp; English.
XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents
XX CC for diagnosing a bacterial disease, as components of antibacterial
XX CC vaccines, as targets for antibacterial drugs, to detect the presence of
XX CC A. baumannii and other Acinetobacter species in a sample, in screening
XX CC compounds for the ability to interfere with the A. baumannii life cycle
XX CC or to inhibit A. baumannii infection, and as biocontrol agents for
XX CC plants. The present sequence represents the amino acid sequence of an A.
XX CC baumannii protein.
XX SQ Sequence 289 AA;
XX
Query Match 47.1%; Score 49; DB 6; Length 289;
Best Local Similarity 38.1%; Pred. No. 25;
Matches 8; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 3 DNRVLTALGVQALQVROIA 23
DB 178 ENHTWAGGIAOAIKAKHQIA 198

```

RESULT 12

```

ADBI6886
ID ADBI6886 standard; peptide; 23 AA.
XX
XX ADBI6886;
XX DT 20-NOV-2003 (first entry)
XX DE Haemophilus influenzae ORF HI0638 penetrating peptide 1.
XX KM penetrating peptide; epithelial; endothelial; tight junction; diabetes;
XX KM infertility; hormone; vitamin deficiency; neurodegenerative;
XX KM cardiovascular; haematological; endocrine disorder; obesity;
XX KM neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
XX KM osteopathic; cytosolic; nootropic.
XX OS Haemophilus influenzae.
XX PF

```

```

PN W02003066859-A2.
XX
XX PD 14-AUG-2003.
XX PF 07-FEB-2003; 2003WO-1B000968.
XX PR 07-FEB-2002; 2002US-0355396P.
XX PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX PI Ben-Sasson SA, Cohen E;
XX PN WPI; 2003-697452/66.
XX PT New penetrating peptide, useful for preparing a composition for treating
XX PT or preventing e.g. endocrine disorders.
XX PS Claim 2; Page 14; 60pp; English.
XX CC This invention relates to a novel peptide sequences capable of
XX CC translocating across a biological barrier. Furthermore, it refers to
XX CC methods that use these peptides to facilitate penetration of a
XX CC biologically active effector molecule such as a drug or other therapeutic
XX CC agent across biological barriers e.g. epithelial or endothelial cells
XX CC sealed by tight junctions. This peptide is derived from a bacterial
XX CC toxin, an integral membrane or extracellular protein and can comprise an
XX CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX CC or enzyme. The effector molecule, however, can comprise for example
XX CC insulin, gonadotropin, erythropoietin, granulocyte/macocyte colony
XX CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
XX CC factors. The penetrating peptide is useful for the treatment of various
XX CC conditions including diabetes, infertility, hormone and vitamin
XX CC deficiencies, neurodegenerative, cardiovascular, haematological and
XX CC endocrine disorders, as well as obesity and neoplastic disease.
XX CC Accordingly, the peptides of this invention can be used in compositions
XX CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
XX CC cytosolic or nootropic activities. This peptide is from ORF HI0638 of
XX CC Haemophilus influenzae and is penetrating peptide 1 of the invention.
XX SQ Sequence 23 AA;
XX
Query Match 46.2%; Score 48; DB 6; Length 23;
Best Local Similarity 61.1%; Pred. No. 1.9;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 6 VTLAAGVQSAKVHQA 23
DB 6 VTLAAGVQSAKVHQA 23

```

RESULT 13

```

ADBI6912
ID ADBI6912 standard; peptide; 23 AA.
XX
XX ADBI6912;
XX DT 20-NOV-2003 (first entry)
XX DE Haemophilus influenzae ORF HI0638 penetrating peptide 34.
XX KM penetrating peptide; epithelial; endothelial; tight junction; diabetes;
XX KM infertility; hormone; vitamin deficiency; neurodegenerative;
XX KM cardiovascular; haematological; endocrine disorder; obesity;
XX KM neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
XX KM osteopathic; cytosolic; nootropic.
XX OS Haemophilus influenzae.
XX PN W02003066859-A2.
XX PD 14-AUG-2003.
XX PF 07-FEB-2003; 2003WO-1B000968.

```

XX 07-FEB-2002; 2002US-0355396P.
PR (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX Ben-Sasson SA, Cohen E;
XX MPI; 2003-697452/66.
XX New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
XX Claim 2; Page 15; 60pp; English.
XX This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
CC cytostatic or neurotrophic activities. This peptide is from ORF HI0638 of
CC Haemophilus influenzae and is penetrating peptide 34 of the invention.
XX
SQ Sequence 23 AA;
Query Match 46.2%; Score 48; DB 6; Length 23;
Best Local Similarity 61.1%; Pred. No. 1.9;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 6 VTLAGVQALQOVQRIA 23
6 VTLAGVCGSARLVHQLA 23
Db
RESULT 14
AEB08218 ID AEB08218 standard; peptide; 23 AA.
XX
AC AEB08218;
XX
DT 25-AUG-2005 (first entry)
DE Haemophilus influenzae ORF HI0638 penetrating peptide 1, SEQ ID NO: 1.
XX
XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
KM antidiabetic; endocrine disease; gastrointestinal disease;
KM metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
KM degenerative; musculoskeletal disease; ocular disease; ophthalmological;
KM neurodegenerative disease; neuroprotective; Alzheimer's disease;
KM neurotrophic; neurological disease; parkinson's disease; antiparkinsonian;
KM dementia; multiple sclerosis; immune disorder; Huntington's chorea;
KM anticonvulsant; genetic disorder; cardiovascular disease;
KM cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
KM coronary artery disease; cardiac; vasotropic; obesity; anorectic;
KM nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
KM genitourinary disease; hematological disease; anitaneimic; anemia;
KM autoimmune disease; immunosuppressive; immune deficiency;
KM immunostimulant; infectious disease; antimicrobial; infection;
KM erectile dysfunction; andrology; major depressive disorder;
KM antidepressant; psychiatric disorder; pain; analgesic;
KM bacterial infection; antibacterial; viral infection; virocidic;
KM fungal infection; fungicide; parasitic infection; antiparasitic;

KM renal failure; antifertility; antirheumatic; cytostatic;
KM antinflammatory; hepatotropic; hepatitis B virus infection.
XX
OS Haemophilus influenzae.
XX
PN US2005136103-A1.
XX
XX 23-JUN-2005.
PD
XX 16-SEP-2004; 2004US-00942300.
PF
XX 17-SEP-2003; 2003US-00664989.
PR 17-SEP-2003; 2003US-00665184.
PR 17-SEP-2003; 2003US-0503615P.
XX
PA (BENS/) BEN-SASSON S A.
PA (COHE/) COHEN E.
PI Ben-Sasson SA, Cohen E;
PI MPI; 2005-444089/45.
DR
XX Composition used for translocating effectors across barrier such as
PT epithelial cells during treatment of e.g. endocrine disorders comprises
PT effector sequentially coupled with counter ion and hydrophobic agent.
XX
XX Claim 46; SEQ ID NO 1; 59pp; English.
XX
XX The present invention relates to a pharmaceutical composition of
CC penetrating peptides for trans epithelial delivery of effector. The
CC invention comprises the effector sequentially coupled with a counter ion
CC and at least one hydrophobic agent, where the effector is selectively
CC encapsulated into a complex. The invention is useful for translocating
CC effectors across a biological barrier such as epithelial cells and
CC endothelial cells during treatment and prevention of disease or
CC pathological conditions (including endocrine disorders, diabetes,
CC infertility, hormone deficiencies, osteoporosis, ophthalmological
CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
CC Parkinson's disease, multiple sclerosis, Huntington's disease,
CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
CC coagulable states, coronary disease, cerebrovascular events, metabolic
CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC hematological disorders, anemia of different entities, immunologic and
CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
CC infectious diseases, viral infections, bacterial infections, fungal
CC infections, parasitic infections, neoplastic diseases, multi-factorial
CC disorders, impotence, chronic pain, depression, different fibrosis states
CC and short stature) and for mucosal vaccination against anthrax and
CC hepatitis B. The present sequence is the Haemophilus influenzae ORF
CC HI0638 penetrating peptide. This sequence is used in the effective
CC translocation of aminoglycoside antibiotics and antifungal agents across
CC an epithelial barrier.
XX
SQ Sequence 23 AA;
Query Match 46.2%; Score 48; DB 9; Length 23;
Best Local Similarity 61.1%; Pred. No. 1.9;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 6 VTLAGVQALQOVQRIA 23
6 VTLAGVCGSARLVHQLA 23
Db
RESULT 15
AEB08246 ID AEB08246 standard; peptide; 23 AA.
XX
AC AEB08246;
XX
DT 25-AUG-2005 (first entry)
DE Haemophilus influenzae ORF HI0638 penetrating peptide 34, SEQ ID NO: 29.

XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimers disease;
 KW neurologic; neurological disease; parkinsons disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antiameic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Haemophilus influenzae.
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 PD 16-SEP-2004; 2004US-00942300.
 XX
 PF 17-SEP-2003; 2003US-00664989.
 XX
 PR 17-SEP-2003; 2003US-00665184.
 XX
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 XX
 DR WPI; 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 29; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Haemophilus influenzae ORF
 CC H10638 penetrating peptide. This sequence is used in the effective
 CC translocation of aminoglycoside antibiotics and antifungal agents across
 CC an epithelial barrier.
 XX
 XX Sequence 23 AA;
 SQ

Query Match 46.2%; Score 48; DB 9; Length 23;
 Best Local Similarity 61.1%; Pred. No. 1.9;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 6 VIALAGVVALDQVRQIA 23
 ||||| : : : :
 Db 6 VIALAGVCSARLVHQLA 23

Search completed: January 23, 2006, 09:20:52
 Job time : 71.2 secs

This Page Blank (except for)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-7

Perfect score: 104

Sequence: 1 LIDNRVTLAGVVALQOVROIA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	104	100.0	204	conserved hypochet
2	50	48.1	2	hypothetical prote
3	49	47.1	2	conserved hypochet
4	49	47.1	2	probable phosphori
5	48	46.2	2	hypothetical prote
6	48	46.2	2	phosphoprotein pho
7	48	46.2	2	folyl-polyglutamat
8	48	46.2	2	phosphoribosylamin
9	47	45.2	2	nicotinate-nucleot
10	47	45.2	2	hypothetical prote
11	46	44.2	2	probable nicotinat
12	46	44.2	2	nicotinate-nucleot
13	46	44.2	2	phosphoprotein pho
14	46	44.2	2	hypothetical prote
15	46	44.2	2	conserved hypochet
16	46	44.2	2	polyketide synthas
17	45	43.3	2	hypothetical prote
18	45	43.3	2	glutamyl-tRNA (Gln)
19	44	42.3	2	conserved hypochet
20	44	42.3	2	yccC protein - Bac
21	44	42.3	2	hypothetical prote
22	44	42.3	2	hypothetical prote
23	44	42.3	2	probable ABC-type
24	44	42.3	2	phosphoprotein pho
25	44	42.3	2	PTS system, fructo
26	43	41.3	2	bis(5'-nucleosyl)-
27	43	41.3	2	hypothetical prote
28	43	41.3	2	fix23-3 protein -
29	42.5	40.9	2	probable general s

30	42	40.4	275	2	B55224	hypothetical prote
31	42	40.4	305	1	A47727	phosphoprotein pho
32	42	40.4	461	2	A31858	ribonuclease-angio
33	42	40.4	490	2	B83508	transcription regu
34	42	40.4	831	2	E70620	probable pter prot
35	42	40.4	840	2	H86429	hypothetical prote
36	42	40.4	852	2	T01364	homeodomain transc
37	41	39.4	127	2	G97716	hypothetical prote
38	41	39.4	184	2	G70648	hypothetical prote
39	41	39.4	215	2	AB0647	conserved hypochet
40	41	39.4	267	2	C83979	myo-inositol-1-ol
41	41	39.4	282	2	G83080	nicotinate-nucleot
42	41	39.4	284	2	C64527	M protein - Helico
43	41	39.4	289	2	G69663	nicotinate-nucleot
44	41	39.4	296	2	F82077	nicotinate-nucleot
45	41	39.4	297	2	T46864	nicotinate-nucleot

ALIGNMENTS

RESULT 1
B82680
conserved hypothetical protein Xf1439 [imported] - Xylella fastidiosa (strain 9a5c)
C/Species: Xylella fastidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: B82680
R/Anonymous: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A82515; PMID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: B82680
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-204 <SIM>
A/Cross-references: UNIPROT:Q9PDB0; UNIPARC:UPI00000C2723; GB:AE003974; GB:AE003849; NID
A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reineck, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre,
as-Neto, R.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitzajima, J.P.; Krieger, J.B.; Kurame, E.B.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.;
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Teshako, M.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A/Reference number: A59328
A/Contents: annotation
C/Genetics:
A/Genes: Xf1439

Query Match 100.0%; Score 104; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIDNRVTLAGVVALQOVROIA 23
Db 4 LIDNRVTLAGVVALQOVROIA 26

RESULT 2
F83597
hypothetical protein PA0391 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: F83597
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yun, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A/Reference number: AB2950; MUID:20437337; PMID:10984043
 A/Accession: F83597
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-655 <STO-
 A/Cross-references: UNIPROT:Q916A4; UNIPARC:UPI000000503B; GB:AE004476; GB:AE004091; NID
 A/Experimental source: strain PA01
 C/Genetics:
 A/Gene: PA0391

Query Match 48.1%; Score 50; DB 2; Length 655;
 Best Local Similarity 47.6%; Pred. No. 9.4;
 Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 LIDNRVLAGVVALQOVRO 21
 |::|||::|||::|||
 DB 588 LLESHMLRLKGLAQERQVRO 608

RESULT 3
 B82237
 conserved hypothetical protein VC1127 [imported] - *Vibrio cholerae* (strain N16961 serogr
 C/Species: *Vibrio cholerae*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: B82237
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 Chardonn, D.; Esmolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, F
 I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A/Reference number: AB2035; MUID:20406833; PMID:10952301
 A/Accession: B82237
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-205 <HEI-
 A/Cross-references: UNIPROT:Q9KX9; UNIPARC:UPI000000C2EB4; GB:AE004193; GB:AE003852; NID
 C/Genetics:
 A/Experimental source: serogroup O1, strain N16961; biotype El Tor
 A/Gene: VC1127
 A/Map position: 1

Query Match 47.1%; Score 49; DB 2; Length 205;
 Best Local Similarity 45.5%; Pred. No. 4.1;
 Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 IDNRVLAGVVALQOVRO 23
 |::|||::|||::|||
 DB 5 IYDRITAFRGICQVALVQOVA 26

RESULT 4
 H71138
 probable phosphoribosylamide-glycine ligase - *Pyrococcus horikoshii*
 C/Species: *Pyrococcus horikoshii*
 C/Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Oct-2004
 C/Accession: H71138
 R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kusuda, N.; Ogunchi
 DNA Res. 5, 55-76, 1998
 A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A/Reference number: A71000; MUID:98344137; PMID:9679194
 A/Accession: H71138
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-438 <KAW-
 A/Cross-references: UNIPROT:O58061; UNIPARC:UPI0000132A3A; GB:AP000002; NID:g3236129; PI
 A/Experimental source: strain OT3
 A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C/Genetics:
 A/Gene: PH0323
 C/Superfamily: phosphoribosylamine-glycine ligase

F/2-416/Domain: phosphoribosylamine-glycine ligase homology <PGL>

Query Match 47.1%; Score 49; DB 2; Length 438;
 Best Local Similarity 39.1%; Pred. No. 8.9;
 Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 LIDNRVLAGVVALQOVRO 23
 |::|||::|||::|||
 DB 374 LLSGRALAIYGVADSLERARIA 396

RESULT 5
 I64155
 hypothetical protein H10638 - *Haemophilus influenzae* (strain Rd KW20)
 C/Species: *Haemophilus influenzae*
 C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
 C/Accession: I64155
 R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995
 A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A/Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 A/Reference number: A64000; MUID:95350630; PMID:7542800
 A/Accession: I64155
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-205 <TIGR-
 A/Cross-references: UNIPROT:P44796; UNIPARC:UPI000013A674; GB:U32747; GB:L42023; NID:g15
 A/Note: best homolog was a hypothetical protein from *Escherichia coli*

Query Match 46.2%; Score 48; DB 2; Length 205;
 Best Local Similarity 61.1%; Pred. No. 5.8;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 VIALAGVVALQOVRO 23
 |||||::|||::|||
 DB 8 VIALAGVVALQOVRO 25

RESULT 6
 S20882
 phosphoprotein phosphatase (EC 3.1.3.16) 1 catalytic chain (clone TOPP1) - *Arabidopsis t*
 N/Alternate names: serine/threonine phosphatase type 1
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)
 C/Date: 16-Sep-1992 #sequence_revision 16-Feb-2001 #text_change 15-Mar-2004
 C/Accession: S20882; S31085; S49037; H84695
 R/Nitschke, K.; Fleig, U.; Schell, J.; Palme, K.
 EMBO J. 11, 1327-1333, 1992
 A>Title: Complementatation of the ca dis2-11 cell cycle mutant of *Schizosaccharomyces pombe*
 A/Reference number: S20882; MUID:92224871; PMID:1314161
 A/Accession: S20882
 A/Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-318 <NIT>
 A/Cross-references: UNIPARC:UPI0000175977; EMBL:X64328; NID:g16430; PIDN:CAA45611.1; PID
 A/Experimental source: clone PPI
 R/Smith, R.D.; Walker, J.C.
 Plant Mol. Biol. 21, 307-316, 1993
 A/Title: Expression of multiple type 1 phosphoprotein phosphatases in *Arabidopsis thalia*
 A/Reference number: S31085; MUID:93144705; PMID:1678768
 A/Accession: S31085
 A/Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-290, 'R', 292-318 <SM11>
 A/Cross-references: UNIPARC:UPI0000175978; EMBL:M93408
 A/Experimental source: clone TOPP1
 R/Smith, R.D.; Walker, J.C.
 submitted to the EMBL Data Library, May 1992
 A/Reference number: S49037
 A/Accession: S49037
 A/Molecule type: mRNA
 A/Residues: 1-318 <SM12>

A/Cross-references: UNIPARC:UPI0000175977; EMBL:M93408; NID:g166571; PIDD:AAA32723.1; P
 A/Experimental source: Clone TOPPI
 R/Lin. X. J. Kaul, S. J. Rounsley, S. D. J. Shea, T. P. J. Benito, M. I. J. Town, C. D. J. Fujii, C. Y. J.
 M. J. Koo, H. J. Moffat, K. S. J. Cronin, L. A. J. Shen, M. J. Vanaken, S. E. J. Umayam, L. J. Tallon, L.
 eues, D. J. Nierman, W. C. J. White, O. J. Eiten, J. A. J. Salberg, S. L. J. Frazer, C. M. J. Venter, J.
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: H84695
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-318 <STO>
 A/Cross-references: UNIPARC:UPI0000175977; GB:AE002093; NID:g3960395; PIDD:AA095198.1; C
 C/Genetics:
 A/Map position: 2
 C/Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phos
 C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-speci
 P/3-301/Domain: phosphoprotein phosphatase homology <PP>
 P/71-139/Domain: phosphoesterase core homology <PPC>
 P/77-79,105/Binding site: iron (Asp, His, Asp) #status predicted
 P/105,137,186,260/Binding site: zinc (Asp, Asn, His, His) #status predicted
 P/108,138,284/Active site: Asp, His, Tyr #status predicted
 P/109,233/Binding site: substrate phosphate (Arg) #status predicted

Query Match 46.2%; Score 48; DB 2; Length 318;
 Best Local Similarity 40.7%; Pred. No. 9.2;
 Matches 11; Conservative 7; Mismatches 5; Indels 4; Gaps 1;

OY 1 LDNRVTLAAGV---QALQVROIA 23
 DB 177 LDNRVTLAAGV---QALQVROIA 203

RESULT 7
 875037
 folyl-polyglutamate synthetase - *Synechocystis* sp. (strain PCC 6803)
 N/Alternate names: protein all1612
 C/Species: *Synechocystis* sp.
 A/Variety: PCC 6803
 C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C/Accession: S75037
 R/Keneo, T. J. Sato, S. J. Kozani, H. J. Tanaka, A. J. Asamizu, E. J. Nakamura, Y. J. Miyajima, N. J.
 o, K. J. Okumura, S. J. Shimpo, S. J. Takeuchi, C. J. Wada, T. J. Watanabe, A. J. Yamada, M. J. Yasuda
 DNA Res. 3, 109-136, 1996
 A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 sp.
 A/Reference number: S74322; MUID:97061201; PMID:8905231
 A/Accession: S75037
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-428 <KAN>
 A/Cross-references: UNIPROT:P73842; UNIPARC:UPI00000034F0; EMBL:D90910; GB:AB001339; NIT
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C/Genetics:
 A/Gene: folC
 C/Superfamily: folylpolyglutamate synthase

Query Match 46.2%; Score 48; DB 2; Length 428;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 LDNRVTLAAGV---QALQVROIA 23
 DB 82 LDNRVTLAAGV---QALQVROIA 103

RESULT 8
 875012
 phosphoribosylamine-glycine ligase (purd) PAB1271 - *Pyrococcus abyssi* (strain Orsay)
 C/Species: *Pyrococcus abyssi*
 C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C/Accession: G75012

R/anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A/Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome str
 A/Reference number: A75001
 A/Accession: G75012
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-438 <KAN>
 A/Cross-references: UNIPROT:Q9U711; UNIPARC:UPI0000132A38; GB:AJ248288; GB:AL096636; NIT
 A/Experimental source: strain Orsay
 C/Genetics:
 A/Gene: PAB1271
 C/Superfamily: phosphoribosylamine-glycine ligase; phosphoribosylamine-glycine ligase h
 P/2-416/Domain: phosphoribosylamine-glycine ligase homology <PGL>

Query Match 46.2%; Score 48; DB 2; Length 438;
 Best Local Similarity 39.1%; Pred. No. 13;
 Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 1 LDNRVTLAAGV---QALQVROIA 23
 DB 374 LDNRVTLAAGV---QALQVROIA 396

RESULT 9
 872696
 nicotinate-nucleotide diphosphorylase (carboxylating) (EC 2.4.2.19) nadc - *Mycobacteriu*
 N/Alternate names: Lep1170_C1_168 protein; quinolinate phosphoribosyltransferase nadc
 C/Species: *Mycobacterium leprae*
 C/Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C/Accession: S72696
 R/Smith, D. R. J. Robinson, K. J.
 submitted to the EMBL Data Library, November 1993
 A/Description: *Mycobacterium leprae* cosmid B1170.
 A/Reference number: S72693
 A/Accession: S72696
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-284 <SMI>
 A/Cross-references: UNIPROT:P46714; UNIPARC:UPI000012PC99; EMBL:U00010; NID:g466780; PI
 A/Genetics:
 C/Superfamily: nicotinate-nucleotide pyrophosphorylase (carboxylating)
 C/Keywords: glycosyltransferase; pentosyltransferase

Query Match 45.2%; Score 47; DB 2; Length 284;
 Best Local Similarity 52.4%; Pred. No. 12;
 Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 3 DNRVTLAAGV---QALQVROIA 23
 DB 172 DNRVTLAAGV---QALQVROIA 192

RESULT 10
 887062
 hypothetical protein nadc [imported] - *Mycobacterium leprae*
 C/Species: *Mycobacterium leprae*
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C/Accession: E87062
 R/Cole, S. T. J. Bigmeier, K. J. Parkhill, J. J. James, K. D. J. Thomson, N. R. J. Wheeler, P. R. J. H
 eam, M. A. J. Rutherford, K. M. J.
 Nature 409, 1007-1011, 2001
 A/Authors: Rutter, S. J. Seeger, K. J. Simon, S. J. Simmonds, M. J. Skelton, J. J. Squares, R. J. S
 A/Title: Massive gene decay in the leprosy bacillus.
 A/Reference number: A86909; MUID:21128732; PMID:11234002
 A/Accession: E87062
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-286 <STO>
 A/Cross-references: UNIPARC:UPI0000165F43; GB:AL450380; NID:g13093177; PIDD:CA031608.1; C
 C/Genetics:

Best Local Similarity 42.1%; Pred. No. 25;
 Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
 QY 3 DNRVLTLAGVVOALQVRQ 21
 |::|||::|||
 Db 352 DSRIDLTGTASMRQMAQ 370

RESULT 15

G82430
 conserved hypothetical protein VCA0681 [imported] - *Vibrio cholerae* (strain N16961 serog
 C)spec:es: *Vibrio cholerae*
 C/date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/accession: G82430
 R/Heldelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.V.;
 Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A/title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A/reference number: A82035; MUID:20406833; PMID:10952301
 A/accession: G82430
 A/status: preliminary
 A/molecule type: DNA
 A/residues: 1-431 <HIS>
 A/cross-references: UNIPROT:Q9KRL1; UNIPARC:UPI00000C35ED; GB:AE004397; GB:AE003853; NID
 A/experimental source: serogroup O1; strain N16961; biotype El Tor
 C/genetics:
 A/gene: VCA0681
 A/map position: 2

Query Match 44.2%; Score 46; DB 2; Length 431;
 Best Local Similarity 47.4%; Pred. No. 26;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 2 IDNRVLTLAGVVOALQVR 20
 |::|||::|||
 Db 361 LPSRIITADVFOALTQAR 379

Search completed: January 23, 2006, 09:32:52
 Job time : 10.1 secs

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 Seconds
(without alignments)
267.334 Million cell updates/sec

Title: US-10-501-838a-7
Perfect score: 104
Sequence: 1 LIDNRVLAAGVVALQVROIA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	204	Y1439_XYLFA	Q9PDE0 xylella fas
2	104	100.0	204	Y661_XYLPT	Q87DML xylella fas
3	84	80.8	204	Y1963_XANCP	Q8P9A2 xanthomonas
4	84	80.8	204	Y1997_XANAC	Q8P109 xanthomonas
5	84	80.8	205	Q5GZRU_XANOR	Q5GZRU xanthomonas
6	84	80.8	205	Q4U0U6_XANCP	Q4U0U6 xanthomonas
7	68	65.4	207	Q60CA8_METCA	Q60CA8 methylotoc
8	57	54.8	205	Y2634_SHEON	Q8EDV9 shewanella
9	52	50.0	289	Q6PFW3_ACICAD	Q6PFW3 acinetobact
10	51	49.0	302	Q6AZR1_XENLA	Q6AZR1 pseudonema
11	50	48.1	127	Q88J61_PSERK	Q88J61 pseudonema
12	50	48.1	243	Q6FCW3_ACICAD	Q6FCW3 acinetobact
13	50	48.1	267	Q6BGS0_PAPTE	Q6BGS0 parametecium
14	50	48.1	286	Q5YYO6_NOCRA	Q5YYO6 nocardia fa
15	50	48.1	347	Q8KB92_CHLRE	Q8KB92 chlorobium
16	50	48.1	655	Q16A4_PSBAB	Q16A4 pseudomonas
17	49	47.1	205	Y1127_VIBCH	Q9KEX9 vibrio chol
18	49	47.1	288	Q8D177_SYNEL	Q8D177 synecococc
19	49	47.1	438	PUR2_PYRHO	Q830E1 pyrococcus
20	48	46.2	205	Y1125_VIBPA	Q87QD0 vibrio para
21	48	46.2	205	Y638_HABIN	Q4A756 haemophilus
22	48	46.2	205	Q4QMS9_HAB18	Q4QMS9 haemophilus
23	48	46.2	284	Q74001_MYCPA	Q74001 mycobacteri
24	48	46.2	318	P611_AKXTH	P03366 atabadiopsis
25	48	46.2	396	Q6NGW8_CORDI	Q6NGW8 coxynebac
26	48	46.2	428	P73842_SYNY3	P73842 synecocyst
27	48	46.2	438	PUR2_PYRAB	Q9UY71 pyrococcus
28	48	46.2	1070	DNAB2_AZOSE	Q5P444 azarococcus sp
29	47	45.2	144	Q82GAI_STRAW	Q82GAI streptomyce
30	47	45.2	205	Y1342_VIBVY	Q7M1E5 vibrio vuln
31	47	45.2	205	Y2927_VIBVU	Q8DBP4 vibrio vuln

RESULT 1
Y1439_XYLFA STANDARD; PRT; 204 AA.
ID Y1439_XYLFA
AC Q9PDE0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein Xf1439.
GN OrderedLocustNames=Xf1439;
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.B., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Pachican A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohlmeier J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.B., Kuramae E.B., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos B.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira A.G., Pereira H.A., Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA da Rosa V.B., Jr., de Sa R.G., Santelli R.V., Sawaaki H.B.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
RA da Silveira J.F., Silveira M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubaki M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Zeldman J., Zetser J.C.,
RT The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC - SIMILARITY: Belongs to the UPF0274 family.
CC Title Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC EMBL, AB003974; AF04248.1; -; Genomic_DNA.

DR PIR; E82680; E82680.
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 DR Complete proteome; Hypothetical protein.
 SQ SEQUENCE 204 AA; 22245 MW; 5D9DE0B3A8B9A3C8 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 204;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIDNRVTLALAGVQALQGVROIA 23
 DB 4 LIDNRVTLALAGVQALQGVROIA 26

RESULT 2
 Y661_XYLFT STANDARD; PRT; 204 AA.
 AC Q87DM1;
 DT 10-OCT-2003 (Rel. 42, Last Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein P00661.
 GN OrderedLocustNames=PD0661;
 OS Xylella fastidiosa (strain Temecula / ATCC 700964).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 NCBI_TaxID=183190;

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX MEDLINE=22421331; PubMed=12533478;
 RX DOI=10.1128/JB.185.3.1018-1026.2003;
 RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
 Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Neon D.H.,
 Takita M.A.S., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
 Goldmann M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
 Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
 Coutinho L.L., Kimura E.T., Ferro E.S., Harkava R., Kuramae E.B.,
 Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
 Bata G.S., Blanco S.R., Brito M.S., Camavan F.S., Celestino A.V.,
 da Cunha A.F., Penille R.C., Ferro J.A., Formiglieri E.F., Kishi L.T.,
 Leonil S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
 da Sousa A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
 Cliverolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
 Kitzajima J.P.;
 RA "Comparative analyses of the complete genome sequences of Pierce's
 RT disease and citrus variegated chlorosis strains of Xylella
 RT fastidiosa.";
 RT J. Bacteriol. 185:1018-1026(2003).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC EMBL; AE012555; AA028532.1; -; Genomic_DNA.
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 DR Complete proteome; Hypothetical protein.
 SQ SEQUENCE 204 AA; 22232 MW; 5E4B498DBA8D14F3 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 204;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LIDNRVTLALAGVQALQGVROIA 23
 DB 4 LIDNRVTLALAGVQALQGVROIA 26

RESULT 3
 Y1963_XANCP STANDARD; PRT; 204 AA.
 AC Q8P9A2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein XCC1963.
 GN OrderedLocustNames=XCC1963;
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 NCBI_TaxID=3340;
 [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
 RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
 Quagga R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
 Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bartolini M.C.,
 Camargo L.E.A., Camarote G., Camavan F., Cardozo J., Chambergo F.,
 Clapina L.P., Claretill R.M.B., Coutinho L.L., Cursino-Santos J.R.,
 El-Dorri H., Faria J.B., Ferreira A.J.S., Greggio C.C., Gruber A.,
 Ferro M.I.T., Formiglieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 localil B.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 Martins E.C., Meidanis J., Menck C.P.M., Miyaki C.Y., Moon D.H.,
 Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira B.C., Tezsa R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
 RA Setubal J.C., Kitzajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RT Nature 417:459-463(2002).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC EMBL; AE012300; AA041252.1; ALT_INIT; Genomic_DNA.
 DR HAMAP; MF_00695; -; 1
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 DR Complete proteome; Hypothetical protein.
 SQ SEQUENCE 204 AA; 22005 MW; DCSFBF027B22BD47 CRC64;

Query Match 80.8%; Score 84; DB 1; Length 204;
 Best Local Similarity 81.8%; Pred. No. 0.00013;
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 IDNRVTLALAGVQALQGVROIA 23
 DB 5 MDNRVTLALAGVQALQGVROIA 26

RESULT 4
 Y1997_XANAC STANDARD; PRT; 204 AA.
 AC Q8PUD5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein XAC1997.
 GN OrderedLocustNames=XAC1997;
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 NCBI_TaxID=92829;

[1]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=306 / ATCC 13302 / XV 101;
 RC MEDLINE=20222145; PubMed=12024217; DOI=10.1038/117459a;
 RX da Silva A.C.R., Ferro J.A., Reinach F.C., Parah C.S., Furlan L.R.,
 RA Almeida N.F., Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
 RA Camargo L.E.A., Camaroto G., Camaván P., Cardoso J., Chambergo F.,
 RA Ciapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
 RA Bl-Dorcy H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
 RA Ferro M.I.T., Fomighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madela A.M.B.N., Martinez-Roset N.M.,
 RA Martins E.C., Melanda J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spínola L.A.F., Takita M.A., Tamura R.B., Teixeira B.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Trufi D., Teal S.M., White F.F.,
 RA Setubal J.C., Kitzima J.P.,
 RT Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463(2002).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL/ AB011835; AAM36859.1; -; Genomic_DNA.
 DR HAMAP; MP 006955; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 204 AA; 21918 MW; CSD87610B5D0F2CF CRC64;

Query Match 80.8%; Score 84; DB 1; Length 204;
 Best Local Similarity 81.8%; Pred. No. 0.00013;
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 IDNRVTLAGVQAALQOVRRIA 23
 :|:|||||:|||||:|
 DB 5 MDHRVTLAGVQAALQOVRRIA 26

RESULT 5
 Q5GZRO_XANOR
 ID Q5GZRO_XANOR PRELIMINARY; PRT; 205 AA.
 AC Q5GZRO;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 DE OrderedLocustNames=XO02557;
 GN Xanthomonas oryzae (pv. oryzae).
 OS Xanthomonas oryzae (pv. oryzae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 NC NCB1_TaxID=64187;
 OX NCBI_TaxID=64187;
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=KACC10331 / KCO85;
 RX PubMed=15673718; DOI=10.1093/nar/gk1206;
 RA Lee B.-M., Park Y.-U., Park D.-S., Kang H.-W., Kim J.-G., Song B.-S.,
 RA Park I.-C., Yoon U.-H., Han J.-H., Koo B.-S., Lee G.-B., Kim H.,
 RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
 RA Go S.-J.;
 RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
 RT the bacterial blight pathogen of rice."
 RL Nucleic Acids Res 33:577-586(2005).
 DR EMBL; AB013598; AAM75811.1; -; Genomic_DNA.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.

KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 205 AA; 22078 MW; 68FD48926EC1BA5A CRC64;

Query Match 80.8%; Score 84; DB 2; Length 205;
 Best Local Similarity 81.8%; Pred. No. 0.00013;
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 IDNRVTLAGVQAALQOVRRIA 23
 :|:|||||:|||||:|
 DB 6 MDHRVTLAGVQAALQOVRRIA 27

RESULT 6
 Q4UTU6_XANCP
 ID Q4UTU6_XANCP PRELIMINARY; PRT; 205 AA.
 AC Q4UTU6;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 DE ORFNames=XC_2222;
 GN Xanthomonas campestris pv. campestris str. 8004.
 OS Xanthomonas campestris pv. campestris str. 8004.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 NC NCB1_TaxID=314565;
 OX NCBI_TaxID=314565;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=8004;
 RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
 RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
 RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-D., Fu G.,
 RA Chen B.-Z., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
 RA He C.-Z.;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CP000050; AAY49277.1; -; Genomic_DNA.
 KM Hypothetical protein.
 SQ SEQUENCE 205 AA; 22136 MW; 32D9A34CB5C95B1 CRC64;

Query Match 80.8%; Score 84; DB 2; Length 205;
 Best Local Similarity 81.8%; Pred. No. 0.00013;
 Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 IDNRVTLAGVQAALQOVRRIA 23
 :|:|||||:|||||:|
 DB 6 MDHRVTLAGVQAALQOVRRIA 27

RESULT 7
 Q60CA8_METCA
 ID Q60CA8_METCA PRELIMINARY; PRT; 207 AA.
 AC Q60CA8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 DE OrderedLocustNames=MC0A0201;
 GN Methylobacterium capsulatus.
 OS Methylobacterium capsulatus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Methylobacterales;
 OC Methylobacteriaceae; Methylobacterium.
 NC NCB1_TaxID=414;
 OX NCBI_TaxID=414;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Bath / NCIMB 11132;
 RX PubMed=15383840; DOI=10.1371/journal.pbio.0020303;
 RA Ward N.L., Larsen O., Sakwa J., Bruseck L., Khouri H.M., Durkin A.S.,
 RA Dmitrov G., Jiang L., Scanlan D., Kang K.H., Lewis M.R., Nelson K.B.,
 RA Mehe B.A., Wu M., Heidelberg J.P., Paulsen I.T., Fouts D.B.,
 RA Ravel J., Tettelin H., Ren Q., Read T.D., Debey R.T., Seshadri R.,
 RA Salzberg S.L., Jensen H.B., Birkeland N.K., Nelson W.C., Dodson R.J.,
 RA Grindhaug S.H., Holt I.E., Rikhammer I., Jonsen I., Vanaken S.,
 RA Unterbach T.R., Feldblyum T.V., Fraser C.M., Lillehaug J.R.,
 RA Elesen J.A.;
 RT "Genomic insights into mechanotrophy: the complete genome sequence of

RT Methylococcus capsulatus (Bath)."
 RL Plos Biol. 2:1616-1628(2004).
 DR EMBL: AB017282; AAU90683.1; -; Genomic_DNA.
 DR TIGR: MCA0201; -.
 DR InterPro: IPR007451; DUF489.
 DR Pfam: PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 207 AA; 22996 MW; C75B763DE01AE86 CRC64;

Query Match 65.4%; Score 68; DB 2; Length 207;
 Best Local Similarity 63.6%; Pred. No. 0.04;
 Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 IDNRVIALAGVVALQGVROIA 23
 DB 5 LDNOVIALAGLQAHTLVROIA 26

RESULT 8
 Y2634 SHEON STANDARD; PRT; 205 AA.
 AC Q8BDV9;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein S02634.
 GN OrderedCusNames=S02634;
 OS Shewanella oneidensis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Shewanellaceae; Shewanella.
 OX NCBI_Taxid=70863;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MR-1;
 RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Galdo E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
 RA Clayton R.A., Meyer T., Tsaplin A., Scott J., Beaman M.J.,
 RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
 RA Haft D.H., Kolony J.F., Madupu R., Peterson J.D., Umayam L.A.,
 RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
 RA Lee K., Berry K.J., Lee C., Mueller J., Khoult H.M., Gill J.,
 RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
 RA Venter J.C., Nealeon K.H., Fraser C.M.;
 RA "Genome sequence of the dissimilatory metal ion-reducing bacterium
 RT Shewanella oneidensis";
 RL Nat. Biotechnol. 20:1118-1123(2002).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC EMBL: AB015704; AAN55662.1; -; Genomic_DNA.
 CC TIGR: SO2634; -; 1.
 DR HAMAP: MF_00695; -; 1.
 DR InterPro: IPR007451; DUF489.
 DR Pfam: PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 205 AA; 22793 MW; CFESB0189677CF31 CRC64;

Query Match 54.8%; Score 57; DB 1; Length 205;
 Best Local Similarity 45.5%; Pred. No. 2;
 Matches 10; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

OY 2 IDNRVIALAGVVALQGVROIA 23
 DB 5 LNRRTMAFGIIOAIAQVOYIA 26

RESULT 9

O6FW3 ACIAD
 ID O6FW3 ACIAD PRELIMINARY; PRT; 289 AA.
 AC O6FW3;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Nicotinate-nucleotide pyrophosphorylase (Quinolinate
 DE phosphoribosyltransferase) (EC 2.4.2.19).
 GN Name=nadC; OrderedLocNames=ACIAD0062;
 OS Acinetobacter sp. (strain ADP1).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Acinetobacter.
 OX NCBI_Taxid=629377;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ADP1;
 RX PubMed=15514110; DOI=10.1093/nar/gkh910;
 RA Barbe V., Vallonet D., Fonknechten N., Kreimeyer A., Ozas S.,
 RA Labarre L., Cruveiller S., Robert C., Duprat S., Winkler P.,
 RA Ornston L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;
 RT ADP1, a versatile and naturally transformation competent bacterium";
 RL Nucleic Acids Res. 32:5766-5779(2004).
 DR EMBL: CR53861; CAG6704.1; -; Genomic_DNA.
 DR GO: GO:0004514; F:nicotinate-nucleotide diphosphorylase (carb. . .; IEA.
 DR GO: GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR GO: GO:0009435; P:NAD biosynthesis; IEA.
 DR InterPro: IPR004393; NadC.
 DR InterPro: IPR002638; Q_phepho_trans.
 DR Pfam: PF01729; QRPase_C; 1.
 DR Pfam: PF02749; QRPase_N; 1.
 DR ProDom: PD003988; Q_phepho_trans; 1.
 DR TIGRFAMs: TIGR00078; nadC; 1.
 KW Complete proteome; Glycosyltransferase; Transferase.
 SQ SEQUENCE 289 AA; 31538 MW; D005D9F649D331C2 CRC64;

Query Match 50.0%; Score 52; DB 2; Length 289;
 Best Local Similarity 38.1%; Pred. No. 17;
 Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

OY 3 DNRVIALAGVVALQGVROIA 23
 DB 178 ENHIMAGGIAOATEKHAQIA 198

RESULT 10
 Q6AZRI XENIA
 ID Q6AZRI XENIA PRELIMINARY; PRT; 302 AA.
 AC Q6AZRI;
 DT 25-OCT-2004 (TRENBLrel. 28, Created)
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
 DE Qpr-prot protein.
 GN Name=qpr-prot;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_Taxid=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buecaw K.H., Scheffer C.F., Bat N.K.,
 RA Hopkin R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stopleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Urdin T.B., Teshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.W.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [3]
 RA NUCLEOTIDE SEQUENCE.
 RP TISSUE=Embryo;
 RC
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC077457; AAH77457.1; -; mRNA.
 DR GO; GO:0004514; P:nicotinate-nucleotide diphosphorylase (carb. . .; IEA.
 DR GO; GO:0004935; P:NAD biosynthesis; IEA.
 DR InterPro; IPR004393; NADc.
 DR InterPro; IPR002638; Q_dphpho_trans.
 DR Pfam; PF01729; QRPase_N; 1.
 DR Pfam; PF02749; QRPase_C; 1.
 DR ProDom; PD003988; Q_dphpho_trans; 1.
 DR TIGRPFAM; TIGR0078; nadC; 1.
 SQ SEQUENCE 302 AA; 32176 MW; 1494C22406FB2A93 CRC64;

Query Match 49.0%; Score 51; DB 2; Length 302;
 Best Local Similarity 52.2%; Pred. No. 25;
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 LDNRVTLALGVVQALQVQVQIA 23
 DB 174 LKDNHVMAGVGKALDGVNALLA 196

RESULT 11
 ID Q6B6J1_PSEPK PRELIMINARY; PRT; 127 AA.
 AC Q6B6J1;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Transcriptional regulator, Weir family.
 GN OrderedLocustNames=Pp2788;
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=160488;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22423060; PubMed=12534463;
 RX DOI=10.1046/j.1462-2920.2002.00366.x;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.B., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
 RA Hance I., Chris Lee P., Holtzapfe B.K., Scanlan D., Tien K.,
 RA Koczar Z., Ueberback T.R., Rizzo M., Lee K., Kosack D., Woestl D.,
 RA Medler H., Lauber J., Stjepandic D., Hohnselt J., Straetz M., Helm S.,
 RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tsemmler B.,
 RA Praser C.M.,
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440,"
 RL Environ. Microbiol. 4:799-808(2002).

DR EMBL; AB016784; AAN68396.1; -; Genomic_DNA.
 DR HSSP; P7565; 1006.
 DR TIGR; PP2788; -;
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR000551; HTH_Merr.
 DR Pfam; PF00376; Merr; 1.
 DR PRINTS; PR00040; HTHMERR.
 DR SMART; SM00422; HTH_MERR; 1.
 DR PROSITE; PS50937; HTH_MERR; 2; 1.
 KW Complete proteome; DNA-binding.
 SQ SEQUENCE 127 AA; 14671 MW; FB11E890273C208F CRC64;

Query Match 48.1%; Score 50; DB 2; Length 127;
 Best Local Similarity 61.1%; Pred. No. 15;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 LDNRVTLALGVVQALQV 19
 DB 83 IDNKKQELMGQIQALQOV 100

RESULT 12
 ID Q6FCW3_ACTAD PRELIMINARY; PRT; 243 AA.
 AC Q6FCW3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=ACTAD1220;
 OS Acinetobacter sp. (strain ADP1).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Acinetobacter.
 OC NCBI_TaxID=62977;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ADP1;
 RX PubMed=15514110; DOI=10.1093/nar/gkh910;
 RA Barbe V., Vallenet D., Fonknechten N., Kreimeyer A., Oztas S.,
 RA Labarre V., Cruevelier S., Robert C., Duprat S., Wincker P.,
 RA Ormston L.N., Weissenbach J., Marliere P., Cohen G.N., Médigue C.,
 RT "Unique features revealed by the genome sequence of Acinetobacter sp.
 ADP1, a versatile and naturally transformation competent bacterium,"
 RL Nucleic Acids Res. 32:5766-5779(2004).
 DR EMBL; CR543861; CAG68096.1; -; Genomic_DNA.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 243 AA; 27334 MW; 00B762C261054400 CRC64;

Query Match 48.1%; Score 50; DB 2; Length 243;
 Best Local Similarity 60.0%; Pred. No. 29;
 Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 NRVTALAGVQALQVQVQIA 23
 DB 17 NRATLALGVVQATQTLTMTA 36

RESULT 13
 ID Q6BG50_PARTB PRELIMINARY; PRT; 267 AA.
 AC Q6BG50;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=PTMB.173;
 OS Paramecium tetraurelia.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillula;
 OC Paramecium.

OK NCBI_TaxID=5888;
 [1]
 NP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Stock d4-2;
 RX PubMed:15296759; DOI=10.1016/j.cub.2004.07.029;
 RA Zagulski M., Nowak J.K., Le Mouel A., Nowacki M., Migalski A.,
 RA Gromadka R., Noel B., Blanc I., Dessen P., Winkler P., Koller A.M.,
 RA Cohen J., Meyer B., Sperling L.;
 RT "High Coding Density on the Largest Parametium tetraurelia Somatic
 RT Chromosome";
 RL Curr. Biol. 14:1397-1404(2004).
 [2]
 NP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Stock d4-2;
 RA Nowak J.K., Migalski A., Gromadka R., Zagulski M.;
 RT "Parametium megabase sequencing project";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR548612; CAH03370.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 267 AA; 31808 MW; E8029770304CFB80 CRC64;

Query Match 48.1%; Score 50; DB 2; Length 267;
 Best Local Similarity 52.6%; Pred. No. 31;
 Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 DNRVTLAAGVQALQOVRO 21
 DB 245 DNRVTLAAGVQALQOVRO 263

RESULT 14
 Q5Y06 NOCPA PRELIMINARY; PRT; 286 AA.
 ID Q5Y06
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Putative quinolinate phosphoribosyltransferase.
 GN Name=ndc; OrderedLocustNames=ndc18390;
 OS Nocardiopsis farcinica.
 OC Bacteriia; Actinobacteriia; Actinobacteriales; Actinomycetales;
 OC Corynebacteriinae; Nocardiaceae; Nocardiaceae;
 OC NCBI_TaxID=37329;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IFM 10152;
 RX PubMed:15466710; DOI=10.1073/pnas.0406410101;
 RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
 RA Shiba T., Hattori M.;
 RT "The complete genomic sequence of Nocardiopsis farcinica IFM 10152";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
 DR EMBL; AP006618; BAD56685.1; -; Genomic_DNA.
 DR SMR; Q5Y06; 19-285.
 DR GO; GO:0004514; F:nicotinate-nucleotide diphosphorylase (carb. . .; IEA.
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR GO; GO:0009435; P:NAD biosynthesis; IEA.
 DR InterPro; IPR004393; Ndc.
 DR InterPro; IPR002638; Q_phospho_trans.
 DR Pfam; PF01729; QRPase_C; 1.
 DR Pfam; PF02749; QRPase_N; 1.
 DR ProDom; PD003988; Q_phospho_trans; 1.
 DR TIGRfam; TIGR00078; ndc; 1.
 KW Complete proteome; Glycosyltransferase; Transferase.
 SQ SEQUENCE 286 AA; 29834 MW; 8B1E86D255120514 CRC64;

Query Match 48.1%; Score 50; DB 2; Length 286;
 Best Local Similarity 52.4%; Pred. No. 34;
 Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 DNRVTLAAGVQALQOVRO 23
 DB 174 DNRVTLAAGVQALQOVRO 194

RESULT 15
 O8KB92 CHLTP PRELIMINARY; PRT; 347 AA.
 ID O8KB92
 AC O8KB92;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=CT1897;
 OS Chlorobium tepidum.
 OC Bacteriia; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
 OC Chlorobaculum.
 OK NCBI_TaxID=1097;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 49652 / DSM 12025;
 RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
 RA Dodson R.J., Deboy R.T., Gwin M.L., Nelson W.C., Haft D.H.,
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
 RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
 RA Parksey D.S., Niernan W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
 RA Radune D., Vamathevan J.J., Khouli H.M., White O., Gruber T.M.,
 RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
 RT "The complete genome sequence of Chlorobium tepidum ATCC 49652, a
 RT photosynthetic, anaerobic, green-sulfur bacterium";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
 DR EMBL; AE006470; AAM73116.1; -; Genomic_DNA.
 DR TIGR; CT1897;
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR006685; Mscs channel.
 DR Pfam; PF00924; Msc channel; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 347 AA; 38590 MW; 8B633FBC14DBB398 CRC64;

Query Match 48.1%; Score 50; DB 2; Length 347;
 Best Local Similarity 58.8%; Pred. No. 41;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 VLAAGVQALQOVRO 22
 DB 143 VLAAGVQALQOVRO 159

Search completed: January 23, 2006, 09:31:12
 Job time : 61.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838a-7

Perfect score: 104

Sequence: 1 LIDNRVTLALAGVQALQGVQRIA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications_AA.Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	23	US-10-665-184-7	Sequence 7, Appli
2	104	100.0	23	US-10-942-300-7	Sequence 7, Appli
3	104	100.0	23	US-10-501-838a-7	Sequence 29, Appli
4	104	100.0	204	US-10-665-184-65	Sequence 65, Appli
5	104	100.0	204	US-10-942-300-65	Sequence 59, Appli
6	104	100.0	204	US-10-501-838a-59	Sequence 68245, A
7	50	48.1	127	US-10-282-122a-68245	Sequence 4, Appli
8	49	47.1	23	US-10-665-184-4	Sequence 4, Appli
9	49	47.1	23	US-10-942-300-4	Sequence 4, Appli
10	49	47.1	23	US-10-501-838a-4	Sequence 4, Appli
11	49	47.1	204	US-10-665-184-62	Sequence 62, Appli
12	49	47.1	204	US-10-942-300-62	Sequence 62, Appli
13	49	47.1	204	US-10-501-838a-56	Sequence 56, Appli
14	49	47.1	23	US-10-665-184-1	Sequence 1, Appli
15	48	46.2	23	US-10-665-184-29	Sequence 29, Appli
16	48	46.2	23	US-10-942-300-1	Sequence 1, Appli
17	48	46.2	23	US-10-942-300-29	Sequence 29, Appli
18	48	46.2	23	US-10-501-838a-1	Sequence 1, Appli
19	48	46.2	23	US-10-501-838a-29	Sequence 29, Appli
20	48	46.2	29	US-10-665-184-33	Sequence 33, Appli
21	48	46.2	29	US-10-942-300-33	Sequence 33, Appli
22	48	46.2	30	US-10-501-838a-33	Sequence 33, Appli
23	48	46.2	205	US-10-665-184-59	Sequence 59, Appli
24	48	46.2	205	US-10-942-300-59	Sequence 59, Appli
25	48	46.2	205	US-10-501-838a-53	Sequence 53, Appli
26	48	46.2	396	US-10-282-122a-53722	Sequence 53722, A
27	47	45.2	144	US-10-156-761-11531	Sequence 11531, A

Applicant

28	47	45.2	230	4	US-10-322-281-450	Sequence 450, App
29	47	45.2	308	4	US-10-282-122a-71040	Sequence 71040, A
30	47	45.2	323	4	US-10-724-972a-1958	Sequence 3958, Ap
31	46	44.2	23	4	US-10-665-184-2	Sequence 2, Appli
32	46	44.2	23	5	US-10-942-300-2	Sequence 2, Appli
33	46	44.2	23	5	US-10-501-838a-2	Sequence 2, Appli
34	46	44.2	203	4	US-10-665-184-60	Sequence 60, Appli
35	46	44.2	203	5	US-10-942-300-60	Sequence 60, Appli
36	46	44.2	203	5	US-10-501-838a-54	Sequence 54, Appli
37	46	44.2	369	3	US-09-801-368-268	Sequence 268, App
38	46	44.2	369	3	US-10-369-493-1568	Sequence 1568, Ap
39	46	44.2	2201	5	US-10-732-923-20535	Sequence 20535, A
40	45	43.3	144	4	US-10-424-599-224409	Sequence 224409, A
41	45	43.3	275	4	US-10-424-599-169986	Sequence 169986, A
42	45	43.3	304	4	US-10-264-049-2583	Sequence 2583, Ap
43	45	43.3	314	4	US-10-156-761-12223	Sequence 12223, A
44	45	43.3	623	4	US-10-104-047-2547	Sequence 2547, Ap
45	44	42.3	22	4	US-10-665-184-27	Sequence 27, Appli

ALIGNMENTS

```

RESULT 1
US-10-665-184-7
; Sequence 7, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCES: 24348-501CIP
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Xylella fastidiosa
US-10-665-184-7
Query Match 100.0%; Score 104; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 LIDNRVTLALAGVQALQGVQRIA 23
Db 1 LIDNRVTLALAGVQALQGVQRIA 23

RESULT 2
US-10-942-300-7
; Sequence 7, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCES: 24348-503
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

```

PRIOR FILING DATE: 2003-09-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 7
LENGTH: 23
TYPE: PRT
ORGANISM: Xylella fastidiosa
US-10-942-300-7

Query Match 100.0%; Score 104; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIDNRVTLAAGVQALQQVROIA 23
DB 1 LIDNRVTLAAGVQALQQVROIA 23

RESULT 3

US-10-501-838A-7
Sequence 7, Application US/10501838A
Publication No. US20050213478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT APPLICATION NUMBER: US/10/501,838A
CURRENT FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 7
LENGTH: 23
TYPE: PRT
ORGANISM: Xylella fastidiosa
US-10-501-838A-7

Query Match 100.0%; Score 104; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIDNRVTLAAGVQALQQVROIA 23
DB 1 LIDNRVTLAAGVQALQQVROIA 23

RESULT 4

US-10-665-184-65
Sequence 65, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 65
LENGTH: 204
TYPE: PRT
ORGANISM: Xylella fastidiosa

US-10-665-184-65

Query Match 100.0%; Score 104; DB 4; Length 204;
Best Local Similarity 100.0%; Pred. No. 5.1e-08;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIDNRVTLAAGVQALQQVROIA 23
DB 4 LIDNRVTLAAGVQALQQVROIA 26

RESULT 5

US-10-942-300-65
Sequence 65, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-503
CURRENT APPLICATION NUMBER: US/10/942,300
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 65
LENGTH: 204
TYPE: PRT
ORGANISM: Xylella fastidiosa
US-10-942-300-65

Query Match 100.0%; Score 104; DB 5; Length 204;
Best Local Similarity 100.0%; Pred. No. 5.1e-08;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIDNRVTLAAGVQALQQVROIA 23
DB 4 LIDNRVTLAAGVQALQQVROIA 26

RESULT 6

US-10-501-838A-59
Sequence 59, Application US/10501838A
Publication No. US20050215478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT APPLICATION NUMBER: US/10/501,838A
CURRENT FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 59
LENGTH: 204
TYPE: PRT
ORGANISM: Xylella fastidiosa
US-10-501-838A-59

Query Match 100.0%; Score 104; DB 5; Length 204;
Best Local Similarity 100.0%; Pred. No. 5.1e-08;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LIDNRVIALAGVVALQOVROIA 23
Db 4 LIDNRVIALAGVVALQOVROIA 26

RESULT 7

US-10-282-122A-68245
Sequence 68245, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zykand, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Yamamoto, Robert
APPLICANT: Foreyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282.122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 68245

LENGTH: 127

TYPE: PRT

ORGANISM: Pseudomonas putida

US-10-282-122A-68245

Query Match

Best Local Similarity 48.1%; Score 50; DB 4; Length 127;

Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 IDNRVIALAGVVALQOV 19

Db 83 IDNRVIALAGVVALQOV 100

RESULT 8

US-10-665-184-4

Sequence 4, Application US/10665184

Publication No. US20040146549A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel

APPLICANT: Cohen, Eilat

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

TITLE OF INVENTION: Biological Barrier

FILE REFERENCE: 24348-501CIP

CURRENT APPLICATION NUMBER: US/10/665,184

CURRENT FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: PCT/IB03/00968

PRIOR FILING DATE: 2003-02-07

PRIOR APPLICATION NUMBER: 60/355,396

PRIOR FILING DATE: 2002-02-07

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.2

SEQ ID NO 4

LENGTH: 23

TYPE: PRT

ORGANISM: Vibrio cholerae

US-10-665-184-4

Query Match

Best Local Similarity 47.1%; Score 49; DB 4; Length 23;

Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IDNRVIALAGVVALQOVROIA 23

Db 2 IDNRVIALAGVVALQOVROIA 23

RESULT 9

US-10-942-300-4

Sequence 4, Application US/10942300

Publication No. US20050136103A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel

APPLICANT: Cohen, Eilat

TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-503

CURRENT APPLICATION NUMBER: US/10/942,300

CURRENT FILING DATE: 2004-09-16

PRIOR APPLICATION NUMBER: 10/665,184

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 10/664,989

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 60/503,615

PRIOR FILING DATE: 2003-09-17

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.2

SEQ ID NO 4

LENGTH: 23

TYPE: PRT

ORGANISM: Vibrio cholerae

US-10-942-300-4

Query Match

Best Local Similarity 47.1%; Score 49; DB 5; Length 23;

Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IDNRVIALAGVVALQOVROIA 23

Db 2 IDNRVIALAGVVALQOVROIA 23

RESULT 10

US-10-501-838A-4

Sequence 4, Application US/10501838A

Publication No. US20050215478A1

GENERAL INFORMATION:

APPLICANT: Ben-Sasson, Shmuel A.

APPLICANT: Cohen, Eilat

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across

FILE REFERENCE: 24348-501 NATL

CURRENT APPLICATION NUMBER: US/10/501,838A

CURRENT FILING DATE: 2004-07-19

PRIOR APPLICATION NUMBER: PCT/IB03/00968

```
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-501-838a-4
```

```
Query Match      47.1%; Score 49; DB 5; Length 23;
Best Local Similarity 45.5%; Pred. No. 1.5;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
```

```
OY      2 IDNRVTLAAGVQALQOVROIA 23
Db      2 IYDRTIAFAGICQAVLVQOVA 23
```

```
RESULT 11
US-10-665-184-62
; Sequence 62, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-665-184-62
```

```
Query Match      47.1%; Score 49; DB 4; Length 204;
Best Local Similarity 45.5%; Pred. No. 19;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
```

```
OY      2 IDNRVTLAAGVQALQOVROIA 23
Db      5 IYDRTIAFAGICQAVLVQOVA 26
```

```
RESULT 12
US-10-942-300-62
; Sequence 62, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 62
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-942-300-62
```

```
Query Match      47.1%; Score 49; DB 5; Length 204;
Best Local Similarity 45.5%; Pred. No. 19;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
```

```
OY      2 IDNRVTLAAGVQALQOVROIA 23
Db      5 IYDRTIAFAGICQAVLVQOVA 26
```

```
RESULT 13
US-10-501-838a-56
; Sequence 56, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-501-838a-56
```

```
Query Match      47.1%; Score 49; DB 5; Length 204;
Best Local Similarity 45.5%; Pred. No. 19;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
```

```
OY      2 IDNRVTLAAGVQALQOVROIA 23
Db      5 IYDRTIAFAGICQAVLVQOVA 26
```

```
RESULT 14
US-10-665-184-1
; Sequence 1, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 23
; TYPE: PRT
; ORGANISM: haemophilus influenzae
US-10-665-184-1
```

```
Query Match      46.2%; Score 48; DB 4; Length 23;
```


Best Local Similarity 61.1%; Pred. No. 2.2;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 VLALAGVQALQOVROIA 23
| | | | | : : : | |
DB 6 VLALAGVCQSARLVHQIA 23

RESULT 15

US-10-665-184-29
; Sequence 29, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Binat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-665-184-29

Query Match 46.2%; Score 48; DB 4; Length 23;
Best Local Similarity 61.1%; Pred. No. 2.2;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 VLALAGVQALQOVROIA 23
| | | | | : : : | |
DB 6 VLALAGVCQSARLVHQIA 23

Search completed: January 23, 2006, 11:52:34
Job time : 61 secs

This page blank (user)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-8

Sequence: 1 NLRPIVLAIVIGICAAVFLQQYV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	64.9	235	2 AE3344	Integral membrane
2	62	54.4	1208	2 T05077	hypothetical prote
3	58	50.9	381	2 B86175	protein P19P19.21
4	54	47.4	999	2 T27628	hypothetical prote
5	51	44.7	953	2 S54478	probable membrane
6	50	43.9	120	2 F82942	conserved hypothet
7	50	43.9	912	2 JH0563	metabotropic gluta
8	49.5	43.4	134	2 T05417	hypothetical prote
9	49	43.0	915	2 A49874	metabotropic gluta
10	48.5	42.5	390	1 S17938	xpsf protein - Xan
11	48.5	42.5	390	1 T12057	xpsf protein - Xan
12	48.5	42.5	452	1 D69810	phosphotransferase
13	48	42.1	243	2 AG1338	FMN-containing NAD
14	48	42.1	286	2 B70632	CDPdiacylglycerol-
15	48	42.1	672	2 B84782	probable receptor-
16	48	42.1	908	2 I49142	metabotropic gluta
17	47.5	41.7	94	2 D71498	probable yops/flig
18	47.5	41.7	3951	2 VF1H81	F1 protein - avian
19	47	41.2	216	2 A11484	transcription regu
20	47	41.2	276	2 F75300	hypothetical prote
21	47	41.2	335	2 A13453	transcription regu
22	47	41.2	369	2 G69254	dolichol-P-glucose
23	47	41.2	670	2 G86702	potassium uptake p
24	47	41.2	718	2 S74509	polysubunit leucide
25	46.5	40.8	95	2 F86593	yops translocation
26	46.5	40.8	95	2 B72030	yops translocation
27	46.5	40.8	145	2 H70946	hypothetical prote
28	46.5	40.8	335	2 T29318	hypothetical prote
29	46.5	40.8	605	2 T12400	NADH2 dehydrogenas

30	46	40.4	79	2 B84223	hypothetical prote
31	46	40.4	82	2 B83740	hypothetical prote
32	46	40.4	154	2 A44809	fimbrial protein -
33	46	40.4	155	2 JH0071	alpha-pilin - Mora
34	46	40.4	156	2 A41490	beta precursor -
35	46	40.4	157	2 A24434	prepilin - Moraxel
36	46	40.4	158	2 A55851	type 4 pilin - Elk
37	46	40.4	159	2 A47699	fimbrial protein Q
38	46	40.4	164	2 A42460	hypothetical prote
39	46	40.4	319	2 C44020	probable iron-cont
40	46	40.4	387	2 F82998	hypothetical prote
41	46	40.4	449	2 S76839	transcription fact
42	46	40.4	466	2 A47561	myocyte enhancer-b
43	46	40.4	473	2 A47284	hypothetical prote
44	45.5	39.9	263	2 AG1199	transglycosylase-a
45	45	39.5	90	2 AB2192	

ALIGNMENTS

RESULT 1
AE3344
Integral membrane protein (rhomboid family) [Imported] - Brucella melitensis (strain 16;
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #ext_change 05-Oct-2004
C/Accession: AE3344
R:/DeVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Petra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elser, P.H.; Hagius, S.; O'Callaghan, D.; Leter-
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen Brucella meliten-
A/Reference number: AD3252; PMID:11756688
A/Accession: AE3344
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-235 <KUR>
A/Cross-references: UNIPROT:Q8YH07; UNIPARC:UPI0000057DCD; GB:AE008917; PIDN:AAU51920.1
A/Experimental source: strain 16M
A/Gene: BME10739
A/Map position: 1
C/Superfamily: AARA protein

Query Match 64.9%; Score 74; DB 2; Length 235;
Best Local Similarity 47.8%; Pred. No. 0.0028; Mismatches 4; Indels 0; Gaps 0;
Matches 11; Conservative 8

QY 1 NLRPIVLAIVIGICAAVFLQQYV 23
DB 12 NLRGVIALIGLCVAVYVYNTI 34

RESULT 2
T05077
hypothetical protein T6K21.90 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #ext_change 09-Jul-2004
C/Accession: T05077
R:/Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Bancroft, I.; Mewes,
submitted to the Protein Sequence Database, February 1998
A/Reference number: Z15397
A/Accession: T05077
A/Molecule type: DNA
A/Residues: 1-1208 <BEV>
A/Cross-references: UNIPROT:O49690; UNIPARC:UPI00000A5885; EMBL:AL021869
A/Experimental source: cultivar Columbia; BAC clone T6K21
C/Genetics:
A/Map position: 4
A/Intons: 153/1; 270/1; 423/3; 486/3; 517/3; 588/1; 628/1; 653/3; 671/3; 701/3; 726/2;
F/1049-1100/Domain: RING finger homology <RNN>
Query Match 54.4%; Score 62; DB 2; Length 1208;

Best Local Similarity 54.5%; Pred. No. 0.59;
Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 1 NLPPIVAVIGICAVFLQOY 22
DB 955 NLSPLVIAIFGIATVAPFLAAV 976

RESULT 3

protein P19P19.21 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: B86175
R/Theologer, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Ansen, N.F.; Hughes, B.; Hultar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maitl, R.; Marziani, R.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: B86175
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-381 <STO>
A/Cross-references: UNIPROT:P93823; UNIPARC:UPI000000C320; GB:AE005172; NID:g1903359; P1
C/Genetic8:
A/Gene: P19P19.21
A/Map position: 1

Query Match 50.9%; Score 58; DB 2; Length 381;
Best Local Similarity 54.5%; Pred. No. 0.82;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 1 NLPPIVAVIGICAVFLQOY 22
DB 40 NPLIAIVIGIATVAPFLVY 61

RESULT 4

hypothetical protein ZC506.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T27628
R/Harris, B.
submitted to the EMBL Data Library, December 1994
A/Reference number: Z20395
A/Accession: T27628
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-999 <WII>
A/Cross-references: UNIPROT:Q09630; UNIPARC:UPI000012F07C; EMBL:Z47073; PIDN:CAA87374.1;
A/Experimental source: clone ZC506
C/Genetic8:
A/Gene: CBSP.ZC506.4
A/Map position: X
A/Introns: 32/2; 65/3; 77/3; 113/2; 140/3; 177/2; 240/2; 277/1; 315/2; 439/2; 471/3; 591
C/Superfamily: metabotropic glutamate receptor 4

Query Match 47.4%; Score 54; DB 2; Length 999;
Best Local Similarity 47.8%; Pred. No. 7;
Matches 11; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 1 NLPPIVAVIGICAVFLQOY 23
DB 682 SLVPTIIVAFGIATVFLVIVYV 704

RESULT 5
S54478
probable membrane protein YMR266w - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YMR156.08
C/Species: Saccharomyces cerevisiae
C/Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C/Accession: S54478
R/Lye, G.; Church, C.M.
submitted to the EMBL Data Library, May 1995
A/Reference number: S54014
A/Accession: S54478
A/Molecule type: DNA
A/Residues: 1-953 <LYE>
A/Cross-references: UNIPROT:Q03516; UNIPARC:UPI000013B93B; EMBL:Z49260; NID:g809081; PID
A/Experimental source: strain AB972
C/Genetics:
A/Gene: SCD:RSL1; MIBS:YMR266w
A/Cross-references: SCD:S0004879
A/Map position: 13R
C/Superfamily: Yeast probable membrane protein YOL084w
C/Keywords: transmembrane protein

F.32-48/Domain: transmembrane #status predicted <TM1>
F.106-122/Domain: transmembrane #status predicted <TM2>
F.152-168/Domain: transmembrane #status predicted <TM3>
F.1395-411/Domain: transmembrane #status predicted <TM4>
F.435-451/Domain: transmembrane #status predicted <TM5>
F.545-561/Domain: transmembrane #status predicted <TM6>
F.599-615/Domain: transmembrane #status predicted <TM7>
F.646-662/Domain: transmembrane #status predicted <TM8>
F.668-684/Domain: transmembrane #status predicted <TM9>

Query Match 44.7%; Score 51; DB 2; Length 953;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 4 PIVAVIGICAVFL 18
DB 668 PIVAVIGICAVFLV 682

RESULT 6

F82942
conserved hypothetical U0041 [imported] - Ureaplasma urealyticum
C/Species: Ureaplasma urealyticum
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: F82942
R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Casseil, G.H.
submitted to GenBank, February 2000
A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A/Reference number: A82870
A/Accession: F82942
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-120 <GLA>
A/Cross-references: UNIPARC:UPI0000139172; GB:AE002103; GB:AF222894; NID:g6898977; PIDN:
A/Experimental source: serovar 3; biovar 1
C/Genetics:
A/Gene: U0041
A/Genetic code: SGC3

Query Match 43.9%; Score 50; DB 2; Length 120;
Best Local Similarity 40.0%; Pred. No. 4.2;
Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
QY 4 PIVAVIGICAVFLQOY 23
DB 5 PVIIVIGIFGIFLVIYGI 24

RESULT 7

JH0563
metabotropic glutamate receptor 4 precursor - rat
C/Species: Rattus norvegicus (Norway rat)

C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
 C/Accession: JH0563; 158149
 R/Tanabe, Y.; Masu, M.; Ichii, T.; Shigemoto, R.; Nakanishi, S.
 Neuron 8, 169-179, 1992
 A/Title: A family of metabotropic glutamate receptors.
 A/Reference number: JH0561; MVID:92110002; PMID:1309649
 A/Accession: JH0563
 A/Molecule type: mRNA
 A/Residues: 1-912 <TAN>
 A/Cross-references: UNIPROT:P31423; UNIPARC:UPI000012F080
 A/Experimental source: brain
 R/O'Hara, P.O.; Sheppard, P.O.; Thøgersen, H.; Venezia, D.; Haldeman, B.A.; McGrane, V.;
 Neuron 11, 41-53, 1993
 A/Title: The ligand-binding domain in metabotropic glutamate receptors is related to bac
 A/Reference number: 158149; MVID:93332699; PMID:8338667
 A/Accession: 158149
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-123, 'R', 125-912 <RES>
 A/Cross-references: UNIPARC:UPI0000167NAE; GB:M90518; NID:G205400; PIDN:AAA93190.1; PID:
 C/Comment: This protein is coupled to a G protein and evokes a variety of functions by m
 C/Genetics:
 A/Gene: GLUR4
 C/Superfamily: metabotropic glutamate receptor 4
 C/Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phospho
 P.1-3/Domain: signal sequence #status predicted <SIG>
 P.13-912/Product: metabotropic glutamate receptor 4 #status predicted <MBT>
 P.158-610/Domain: transmembrane #status predicted <TR>
 P.1625-645/Domain: transmembrane #status predicted <TR>
 P.1657-675/Domain: transmembrane #status predicted <TR>
 P.1700-1720/Domain: transmembrane #status predicted <TR>
 P.1751-1772/Domain: transmembrane #status predicted <TR>
 P.1786-807/Domain: transmembrane #status predicted <TR>
 P.822-847/Domain: transmembrane #status predicted <TR>
 P.98,301,454,484,563/Binding site: carboxylate (asn) (covalent) #status predicted
 P.621,689,695,859,870/Binding site: phosphate (ser) (covalent) #status predicted

Query Match 43.9%; Score 50; DB 2; Length 912;
 Best Local Similarity 45.0%; Pred. No. 24;
 Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 4 PIVLAIVIGICAVFLIQYV 23
 DB 591 PLPLAVGVIGATLVVVTVFV 610

RESULT 8
 T05417
 hypothetical protein P28A23.50 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C/Accession: T05417
 R/Bevan, M.; Weichselgartner, M.; Pattmann, B.; Granderath, K.; Dauner, D.; Herzl, A.; N
 submitted to the Protein Sequence Database, October 1998
 A/Reference number: Z15415
 A/Accession: T05417
 A/Molecule type: DNA
 A/Residues: 1-134 <BEV>
 A/Cross-references: UNIPROT:O49486; UNIPARC:UPI0000041555; EMBL:AL021961
 A/Experimental source: cultivar Columbia; BAC clone P28A23
 C/Genetics:
 A/Map position: 4
 A/Introns: 41/1; 100/3
 A/Note: P28A23.50

Query Match 43.4%; Score 49.5; DB 2; Length 134;
 Best Local Similarity 50.0%; Pred. No. 5.5;
 Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 2 LPIVLAIV---IGICAVFLIQ 20
 DB 108 LPIVLAIVTALVGVLAIVFIQ 129

RESULT 9
 A49874
 metabotropic glutamate receptor 7 - rat
 N/Alternate names: metabotropic glutamate receptor mGluR7
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C/Accession: A49874; 157954
 R/Okamoto, N.; Horii, S.; Akazawa, C.; Hayashi, Y.; Shigemoto, R.; Mizuno, N.; Nakanishi,
 J. Biol. Chem. 269, 1231-1236, 1994
 A/Title: Molecular characterization of a new metabotropic glutamate receptor mGluR7 cou
 A/Reference number: A49874; MVID:94117433; PMID:8288585
 A/Accession: A49874
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-915 <RES>
 A/Cross-references: UNIPROT:P35400; UNIPARC:UPI000012F084; GB:D16817; NID:G458728; PIDN
 R/Saugstad, J.A.; Kinzie, J.M.; Mulvihill, E.R.; Segerson, T.P.; Westbrook, G.L.
 Mol. Pharmacol. 45, 367-372, 1994
 A/Title: Cloning and expression of a new member of the L-2-amino-4-phosphonobutyric ac
 A/Reference number: 157954; MVID:94195260; PMID:8145723
 A/Accession: 157954
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-915 <RES>
 A/Cross-references: UNIPARC:UPI000012F084; EMBL:U06832; NID:G459657; PIDN:AAA20655.1; P
 C/Genetics:
 A/Gene: MGLUR7
 C/Superfamily: metabotropic glutamate receptor 4
 C/Keywords: neurotransmitter receptor

Query Match 43.0%; Score 49; DB 2; Length 915;
 Best Local Similarity 35.0%; Pred. No. 34;
 Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 4 PIVLAIVIGICAVFLIQYV 23
 DB 594 PVPLAVGVIGATLVVVTVFV 613

RESULT 10
 S17938
 xpf protein - Xanthomonas campestris pv. campestris
 C/Species: Xanthomonas campestris pv. campestris
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: S17938
 R/Dums, F.; Dow, J.M.; Daniels, M.J.
 Mol. Gen. Genet. 229, 357-364, 1991
 A/Title: Structural characterization of protein secretion genes of the bacterial phytop
 exia.
 A/Reference number: S17937; MVID:92049233; PMID:1944223
 A/Accession: S17938
 A/Molecule type: DNA
 A/Residues: 1-390 <DUM>
 A/Cross-references: UNIPROT:P31744; UNIPARC:UPI0000105DC; EMBL:X59079; NID:G48896; PID
 C/Genetics:
 A/Gene: xpgf
 C/Superfamily: secretion protein xcps
 C/Keywords: transmembrane protein

Query Match 42.5%; Score 48.5; DB 1; Length 390;
 Best Local Similarity 64.7%; Pred. No. 19;
 Matches 11; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 3 PIVLAIVIGICAVFLI 19
 DB 159 PIVLAIVG-CALIFLL 174

RESULT 11
 T12057
 xpf protein - Xanthomonas campestris pv. campestris
 C/Species: Xanthomonas campestris pv. campestris

CjDate: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
CjAccession: T12057
RjHu, N.T.; Hung, M.N.; Wang, K.C.
submitted to the EMBL Data Library, July 1998
A>Description: A functional promoter is located within the coding sequence of a gene, pe
A.Reference number: Z17398
AjAccession: T12057
AjStatus: preliminary; translated from GB/EMBL/DBJ
AjMolecule type: DNA
AjResidues: 1-390 <HUN>
AjCross-references: UNIPROT:P31744, UNIPARC:UPI0000170SD2, EMBL:L02630, NID:g3329822, PI
AjExperimental source: isolate XC1701
CjGenetics:
AjNote: xpsf
CjSuperfamily: secretion protein xcps

Query Match 42.5%; Score 48.5; DB 2; Length 390;
Best Local Similarity 64.7%; Pred. No. 19;
Matches 11; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 3 PPIYAVIGICAAVFL 19
|||:|||||:|||||
DB 159 PPIYAVIGICAAVFL 174

RESULT 12
D69810
phosphotransferase system enzyme II (EC 2.7.1.69) factor II homolog yfif - Bacillus subt
CjSpecies: Bacillus subtilis
CjDate: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
CjAccession: D69810
RjKunst, F.; Ogasaara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
Nature 390, 249-256, 1997
AjAuthors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Galler
leach, J.; Harwood, C.R.; Hensut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Jandino, M.F.
Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinis
AjAuthors: Lauber, J.; Lazarovic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Poll, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, W.; Sadale, Y.; Sato, T.; Scanlon
AjAuthors: Schleich, S.; Schroeder, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Serot
Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Teperera, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
AjAuthors: Yoshikawa, H.F.; Zumbstein, B.; Yoshikawa, H.; Zundin, A.
AjTitle: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
AjReference number: A69580; MUID:198044033; PMID:9384377
AjAccession: D69810
AjStatus: preliminary; nucleic acid sequence not shown; translation not shown
AjMolecule type: DNA
AjResidues: 1-452 <KUN>
AjCross-references: UNIPROT:O34521, UNIPARC:UPI000006009A, GB:Z99108, GB:AL009126, NID:g
AjExperimental source: strain 168
CjGenetics:
AjGene: yfif
CjFunction:
AjDescription: mediates transport of glucose across the cytoplasmic membrane concomitant
ion

CjSuperfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosphot
CjKeywords: phosphoprotein; phosphotransferase; sugar transport system; transmembrane pr
Fj1-452/Domain: phosphotransferase system glucose-specific enzyme II, factor II homology
Fj12-28/Domain: transmembrane #status predicted <TM01>
Fj46-62/Domain: transmembrane #status predicted <TM02>
Fj94-110/Domain: transmembrane #status predicted <TM03>
Fj133-149/Domain: transmembrane #status predicted <TM04>
Fj253-269/Domain: transmembrane #status predicted <TM05>
Fj286-302/Domain: transmembrane #status predicted <TM06>
Fj329-345/Domain: transmembrane #status predicted <TM07>
Fj419-436/Domain: transmembrane #status predicted <TM08>
Fj397/Active site: Cys (phosphocysteine intermediate) #status predicted

Query Match 42.5%; Score 48.5; DB 1; Length 452;
Best Local Similarity 47.6%; Pred. No. 22;

Matches 10; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 4 PPIYAVIGICAAVFLQQYV 23
|||:|||||:|||||
DB 327 PPIYAVIGICAAVFLQQYV 347

RESULT 13
AG1338
FMN-containing NADPH-linked nitro/flavin reductase homolog lmo2111 [imported] - Listeria
CjSpecies: Listeria monocytogenes
CjDate: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
CjAccession: AG1338
RjGlaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Kartz, U.
Science 294, 849-852, 2001
AjAuthors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
AjTitle: Comparative genomics of Listeria species.
AjReference number: AB1077; MUID:21537279; PMID:11679669
AjAccession: AG1338
AjStatus: preliminary
AjMolecule type: DNA
AjResidues: 1-243 <GLA>
AjCross-references: UNIPROT:O8Y5F3, UNIPARC:UPI0000054F73; GB:NC_003210; PIDN:CAD00189.1
AjExperimental source: strain BGD-e
CjGenetics:
AjGene: lmo2111
CjSuperfamily: NADPH-flavin oxidoreductase homolog

Query Match 42.1%; Score 48; DB 2; Length 243;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NLPPYVAVIGICAAV 16
|||:|||||:|||||
DB 150 NLPPYVAVIGICAAV 165

RESULT 14
B70632
CDPdiacylglycerol--serine O-phosphatidyltransferase (EC 2.7.8.8) [similarity] - Mycobacte
CjSpecies: Mycobacterium tuberculosis
CjDate: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004
CjAccession: B70632
RjCole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
AjAuthors: Sgares, R.; Sullivan, J.B.; Taylor, K.; Whitehead, S.; Barrett, B.G.
AjTitle: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
AjReference number: A70500; MUID:98295987; PMID:9634230
AjAccession: B70632
AjStatus: preliminary; nucleic acid sequence not shown; translation not shown
AjMolecule type: DNA
AjResidues: 1-286 <COL>
AjCross-references: UNIPROT:P96282, UNIPARC:UPI0000132817; GB:Z84724; GB:AL123456; NID:g
AjExperimental source: strain H37Rv
CjGenetics:
AjGene: pssA
CjSuperfamily: CDP-diacylglycerol--serine O-phosphatidyltransferase; Bacillus subtilis C
CjKeywords: transferase
Fj38-193/Domain: Bacillus subtilis CDPdiacylglycerol--serine O-phosphatidyltransferase ho

Query Match 42.1%; Score 48; DB 2; Length 286;
Best Local Similarity 40.7%; Pred. No. 17;
Matches 11; Conservative 6; Mismatches 6; Indels 4; Gaps 1;

QY 1 NLPPYVAVIGICAAVFLQQYV 23
|||:|||||:|||||
DB 194 SVPPYVAVIGICAAVFLQQYV 220

RESULT 15

B84782

probable receptor-like protein kinase [Imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 31-Dec-2004

C/Accession: B84782

R/Lin: X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: AB4420; MIMD:20083487; PMID:10617197

A/Accession: B84782

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-672 <STO>

A/Cross-references: UNIPROT:Q9S0J1; UNIPARC:UP100000485FA; GB:AB002093; NID:94581155; PI

C/Genetic8:

A/Gene: At2g36570

A/Map position: 2

C/Superfamily: Receptor-like protein kinase

Query Match 42.1%; Score 48; DB 2; Length 672;

Best local similarity 50.0%; Pred. No. 36;

Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 PPIVAVIGICAVPLQQY 22

DB 268 PGIIVAVIGICAVPLVVSF 287

Search completed: January 23, 2006, 09:32:53
Job time: 10.1 sec

This Page Blank (uspo)

Best Local Similarity 50.0%; Pred. No. 6.9;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 NLPVAVIGICAAVFLQGV 20
Db 96 NAGPFTWAILICIAVFLIQ 115

RESULT 3

US-09-820-809-13
Sequence 13, Application US/09820809
Patent No. 6608176
GENERAL INFORMATION:
APPLICANT: CHAUDHARI, NIRUPA
APPLICANT: ROOPER, STEPHEN D.
TITLE OF INVENTION: TASTE RECEPTOR FOR UMAMI (MONOSODIUM GLUTAMATE) TASTE
FILE REFERENCE: 70373/275576
CURRENT APPLICATION NUMBER: US/09/820,809
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/193,454
PRIOR FILING DATE: 2000-03-31
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 13
LENGTH: 604
TYPE: PRT
ORGANISM: Rattus sp.
US-09-820-809-13

Query Match 43.9%; Score 50; DB 2; Length 604;
Best Local Similarity 45.0%; Pred. No. 18;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 4 PIVAVIGICAAVFLQGV 23
Db 283 PLFLAVGIAATLFFVVTFFV 302

RESULT 4
PCT-US91-09422-19
Sequence 19, Application PC/TUS9109422
GENERAL INFORMATION:
APPLICANT: Mulvihill, Eileen R.
APPLICANT: Hagen, Frederick S.
APPLICANT: Houamed, Khaled M.
APPLICANT: Almers, Wolhard
TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESS: Townsend and Townsend
STREET: One Market Plaza, Stewart Street Tower
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/09422
FILING DATE: 19911212
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/672,007
FILING DATE: 18-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/648,481
FILING DATE: 30-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/626,806
FILING DATE: 12-DEC-1990

ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-6PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 912 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-09422-19

Query Match 43.9%; Score 50; DB 4; Length 912;
Best Local Similarity 45.0%; Pred. No. 28;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 4 PIVAVIGICAAVFLQGV 23
Db 591 PLFLAVGIAATLFFVVTFFV 610

RESULT 5
US-08-481-956A-11
Sequence 11, Application US/08481956A
Patent No. 5824867

GENERAL INFORMATION:
APPLICANT: Coruzzi, Gloria
APPLICANT: Oliveira, Igor
APPLICANT: Lam, Hon-Ming
APPLICANT: Hsieh, Ming-Hsiun
TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESS: Pennie & Edwards
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,956A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 5914-040
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 179 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-481-956A-11

Query Match 43.0%; Score 49; DB 1; Length 179;
Best Local Similarity 35.0%; Pred. No. 7;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 4 PIVAVIGICAAVFLQGV 23

Db 83 PVLAVIGICAVFLQOY 102

RESULT 6

US-08-629-291A-11

Sequence 11, Application US/08629291A

Patent No. 5959174

GENERAL INFORMATION:

APPLICANT: Coruzzi, Gloria

APPLICANT: Oliveira, Igor

APPLICANT: Lam, Hon-Ming

TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESSES:

ADDRESSER: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

ZIP: 10036-2711

COUNTRY: U.S.A.

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 5914-050

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 179 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

Query Match 43.0%; Score 49; DB 1; Length 179;

Best Local Similarity 35.0%; Pred. No. 7;

Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 4 PVLAVIGICAVFLQOY 23

Db 83 PVLAVIGICAVFLQOY 102

RESULT 7

US-08-658-335B-11

Sequence 11, Application US/08658335B

Patent No. 5981703

GENERAL INFORMATION:

APPLICANT: Coruzzi, Gloria

APPLICANT: Oliveira, Igor

APPLICANT: Lam, Hon-Ming

TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESSES:

ADDRESSER: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/658,335B

FILING DATE: 05-JUN-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 5914-052

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 179 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

Query Match 43.0%; Score 49; DB 1; Length 179;

Best Local Similarity 35.0%; Pred. No. 7;

Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 4 PVLAVIGICAVFLQOY 23

Db 83 PVLAVIGICAVFLQOY 102

RESULT 8

US-09-406-640-11

Sequence 11, Application US/09406640

Patent No. 6451546

GENERAL INFORMATION:

APPLICANT: Coruzzi, Gloria

APPLICANT: Oliveira, Igor

APPLICANT: Lam, Hon-Ming

TITLE OF INVENTION: PLANT GLUTAMATE RECEPTORS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESSES:

ADDRESSER: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

SEQUENCE CHARACTERISTICS:
LENGTH: 179 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-406-640-11

Query Match 43.0%; Score 49; DB 2; Length 179;
Best Local Similarity 35.0%; Pred. No. 7;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 4 PIVLAIVIGICAVFLQOYV 23
Db 83 PVLAVLGIATITFVMTPI 102

RESULT 9
US-08-617-785-16
Sequence 16, Application US/08617785E

Patent No. 6228610
GENERAL INFORMATION:
APPLICANT: Flor, Peter J.
APPLICANT: Kuhn, Ranier
APPLICANT: Lindaur, Kristen
APPLICANT: Putner, Irene
APPLICANT: Knopfel, Thomas
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
FILE REFERENCE: 4-19679/A/PCT
CURRENT FILING DATE: 1996-03-19
EARLIER APPLICATION NUMBER: US/08/617,785E
EARLIER FILING DATE: 1994-09-07
EARLIER APPLICATION NUMBER: EPO 9416553.7
EARLIER FILING DATE: 1994-08-19
EARLIER APPLICATION NUMBER: EPO 93810663.0
EARLIER FILING DATE: 1993-09-20
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 210
TYPE: PRT
ORGANISM: Homo sapiens
US-08-617-785-16

Query Match 43.0%; Score 49; DB 2; Length 210;
Best Local Similarity 42.9%; Pred. No. 8.3;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 3 PIVLAIVIGICAVFLQOYV 23
Db 88 PVLAVLGIATITFVMTPI 108

RESULT 10
US-09-817-464-16
Sequence 16, Application US/09817464

Patent No. 6515107
GENERAL INFORMATION:
APPLICANT: Flor, Peter J.
APPLICANT: Kuhn, Ranier
APPLICANT: Lindaur, Kristen
APPLICANT: Putner, Irene
APPLICANT: Knopfel, Thomas
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
FILE REFERENCE: 4-19679/A/PCT
CURRENT FILING DATE: 2001-03-26
EARLIER APPLICATION NUMBER: US/08/617,785
EARLIER FILING DATE: 1996-03-19

EARLIER APPLICATION NUMBER: EPO 9416553.7
EARLIER FILING DATE: 1994-08-19
EARLIER APPLICATION NUMBER: EPO 93810663.0
EARLIER FILING DATE: 1993-09-20
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 210
TYPE: PRT
ORGANISM: Homo sapiens
US-09-817-464-16

Query Match 43.0%; Score 49; DB 2; Length 210;
Best Local Similarity 42.9%; Pred. No. 8.3;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 3 PIVLAIVIGICAVFLQOYV 23
Db 88 PVLAVLGIATITFVMTPI 108

RESULT 11
US-08-617-785-8
Sequence 8, Application US/08617785E

Patent No. 6228610
GENERAL INFORMATION:
APPLICANT: Flor, Peter J.
APPLICANT: Kuhn, Ranier
APPLICANT: Lindaur, Kristen
APPLICANT: Putner, Irene
APPLICANT: Knopfel, Thomas
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
FILE REFERENCE: 4-19679/A/PCT
CURRENT FILING DATE: 1996-03-19
EARLIER APPLICATION NUMBER: EPO 9416553.7
EARLIER FILING DATE: 1994-09-07
EARLIER APPLICATION NUMBER: EPO 93810663.0
EARLIER FILING DATE: 1993-09-20
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 481
TYPE: PRT
ORGANISM: Homo sapiens
US-08-617-785-8

Query Match 43.0%; Score 49; DB 2; Length 481;
Best Local Similarity 35.0%; Pred. No. 20;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 4 PIVLAIVIGICAVFLQOYV 23
Db 169 PVLAVLGIATITFVMTPI 188

RESULT 12
US-09-817-464-8
Sequence 8, Application US/09817464

Patent No. 6515107
GENERAL INFORMATION:
APPLICANT: Flor, Peter J.
APPLICANT: Kuhn, Ranier
APPLICANT: Lindaur, Kristen
APPLICANT: Putner, Irene
APPLICANT: Knopfel, Thomas
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
FILE REFERENCE: 4-19679/A/PCT
CURRENT FILING DATE: 2001-03-26
EARLIER APPLICATION NUMBER: US/09/817,464

```
/ CURRENT FILING DATE: 2001-03-26
/ EARLIER APPLICATION NUMBER: US/08/617,785
/ EARLIER FILING DATE: 1996-03-19
/ EARLIER APPLICATION NUMBER: EPO 9416553.7
/ EARLIER FILING DATE: 1994-08-19
/ EARLIER APPLICATION NUMBER: EPO 93810663.0
/ EARLIER FILING DATE: 1993-09-20
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 8
/ LENGTH: 481
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-817-464-8
```

```
Query Match          43.0%; Score 49; DB 2; Length 481;
Best Local Similarity 35.0%; Pred. No. 20;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 4 PIVAVIGICAAVFLQOYV 23
DB 169 PVFLAMGLIATIFVWATFI 188
```

```
RESULT 13
US-09-134-000C-5470
/ Sequence 5470, Application US/09134000C
/ Patent No. 6617156
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ FILE REFERENCE: 032796-032
/ CURRENT APPLICATION NUMBER: US/09/134,000C
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/055,778
/ PRIOR FILING DATE: 1997-08-15
/ NUMBER OF SEQ ID NOS: 6812
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5470
/ LENGTH: 711
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (21)-(92)
/ OTHER INFORMATION: Amino acids 21, 23, 30, 86 & 92 are Xaa wherein Xaa =
/ OTHER INFORMATION: any amino acid.
/ OTHER INFORMATION: 1
US-09-134-000C-5470
```

```
Query Match          43.0%; Score 49; DB 2; Length 711;
Best Local Similarity 33.3%; Pred. No. 31;
Matches 7; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 3 PIVAVIGICAAVFLQOYV 23
DB 652 PPMVIGIHTFPAIYLSQFI 672
```

```
RESULT 14
US-08-617-785-4
/ Sequence 4, Application US/08617785E
/ Patent No. 6228610
/ GENERAL INFORMATION:
/ APPLICANT: Flor, Peter J.
/ APPLICANT: Kuhn, Rainer
/ APPLICANT: Lindaur, Kristen
/ APPLICANT: Puttner, Irene
/ APPLICANT: Knopfel, Thomas
/ TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
/ TITLE OF INVENTION: HMR6, HMR7) and Related DNA Compounds
/ FILE REFERENCE: 4-19679/A/PCT
```

```
/ CURRENT APPLICATION NUMBER: US/08/617,785E
/ CURRENT FILING DATE: 1996-03-19
/ EARLIER APPLICATION NUMBER: PCT/EP94/02991
/ EARLIER FILING DATE: 1994-09-07
/ EARLIER APPLICATION NUMBER: EPO 9416553.7
/ EARLIER FILING DATE: 1994-08-19
/ EARLIER APPLICATION NUMBER: EPO 93810663.0
/ EARLIER FILING DATE: 1993-09-20
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 4
/ LENGTH: 867
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-08-617-785-4
```

```
Query Match          43.0%; Score 49; DB 2; Length 867;
Best Local Similarity 35.0%; Pred. No. 38;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 4 PIVAVIGICAAVFLQOYV 23
DB 539 PVFLAMGLIATIFVWATFI 558
```

```
RESULT 15
US-09-817-464-4
/ Sequence 4, Application US/09817464
/ Patent No. 6515107
/ GENERAL INFORMATION:
/ APPLICANT: Flor, Peter J.
/ APPLICANT: Kuhn, Rainer
/ APPLICANT: Lindaur, Kristen
/ APPLICANT: Puttner, Irene
/ APPLICANT: Knopfel, Thomas
/ TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
/ FILE REFERENCE: 4-19679/A/PCT
/ CURRENT APPLICATION NUMBER: US/09/817,464
/ CURRENT FILING DATE: 2001-03-26
/ EARLIER APPLICATION NUMBER: US/08/617,785
/ EARLIER FILING DATE: 1996-03-19
/ EARLIER APPLICATION NUMBER: EPO 9416553.7
/ EARLIER FILING DATE: 1994-08-19
/ EARLIER APPLICATION NUMBER: EPO 93810663.0
/ EARLIER FILING DATE: 1993-09-20
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 4
/ LENGTH: 867
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-817-464-4
```

```
Query Match          43.0%; Score 49; DB 2; Length 867;
Best Local Similarity 35.0%; Pred. No. 38;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 4 PIVAVIGICAAVFLQOYV 23
DB 539 PVFLAMGLIATIFVWATFI 558
```

```
Search completed: January 23, 2006, 09:35:48
Job time : 17.6 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:32:56 (Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838A-8

Perfect score: 114
Sequence: 1 NLPPIVAVIGICAVFLQQYV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New.*
1: /cgn2_6/pcdata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/pcdata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/pcdata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/pcdata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/pcdata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/pcdata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/pcdata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/pcdata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	43.0	865	US-11-094-519A-41	Sequence 41, Appl
2	45	39.5	208	US-10-467-657-4924	Sequence 4924, Ap
3	45	39.5	759	US-10-467-657-2722	Sequence 2722, Ap
4	44	38.6	633	US-11-119-683-3	Sequence 3, Appl
5	43	37.7	98	US-10-467-657-842	Sequence 842, App
6	43	37.7	170	US-11-052-554A-238	Sequence 238, App
7	42.5	37.3	76	US-09-978-360A-721	Sequence 721, App
8	42.5	37.3	92	US-09-978-360A-723	Sequence 723, App
9	42	36.8	144	US-10-821-234-1255	Sequence 1255, App
10	42	36.8	204	US-10-858-730-95	Sequence 95, Appl
11	42	36.8	276	US-10-467-657-2200	Sequence 2200, Ap
12	42	36.8	783	US-11-082-389-354	Sequence 354, App
13	41.5	36.4	315	US-10-524-647-25	Sequence 25, Appl
14	41.5	36.4	329	US-10-524-647-23	Sequence 23, Appl
15	41.5	36.4	329	US-10-524-647-80	Sequence 80, Appl
16	41.5	36.4	329	US-10-524-647-27	Sequence 27, Appl
17	40	35.1	257	US-11-093-621-18	Sequence 18, Appl
18	40	35.1	339	US-10-793-626-1694	Sequence 1694, Ap
19	40	35.1	340	US-10-467-657-1010	Sequence 1010, Ap
20	40	35.1	340	US-10-467-657-5288	Sequence 5288, Ap
21	40	35.1	351	US-10-467-657-994	Sequence 994, App
22	40	35.1	410	US-10-467-657-4094	Sequence 4094, App
23	40	35.1	430	US-10-467-657-7680	Sequence 7680, App
24	40	35.1	1436	US-11-052-554A-140	Sequence 140, App
25	39	34.2	232	US-10-467-962B-87	Sequence 87, Appl

26	39	34.2	550	7	US-11-043-889-5	Sequence 5, Appl
27	39	34.2	858	7	US-11-077-550-22	Sequence 22, Appl
28	39	34.2	860	7	US-11-077-550-175	Sequence 175, App
29	39	34.2	862	7	US-11-077-550-94	Sequence 94, Appl
30	39	34.2	862	7	US-11-077-550-171	Sequence 171, App
31	39	34.2	862	7	US-11-077-550-173	Sequence 173, App
32	39	34.2	864	7	US-11-077-550-102	Sequence 102, App
33	39	34.2	865	7	US-11-077-550-100	Sequence 100, App
34	39	34.2	866	7	US-11-077-550-88	Sequence 88, Appl
35	39	34.2	866	7	US-11-077-550-104	Sequence 104, App
36	39	34.2	867	7	US-11-077-550-80	Sequence 80, Appl
37	39	34.2	867	7	US-11-077-550-96	Sequence 96, Appl
38	39	34.2	867	7	US-11-077-550-98	Sequence 98, Appl
39	39	34.2	870	7	US-11-077-550-92	Sequence 92, Appl
40	39	34.2	871	7	US-11-077-550-84	Sequence 84, Appl
41	39	34.2	871	7	US-11-077-550-86	Sequence 86, Appl
42	39	34.2	871	7	US-11-077-550-90	Sequence 90, Appl
43	39	34.2	876	7	US-11-077-550-82	Sequence 82, Appl
44	39	34.2	876	7	US-11-077-550-106	Sequence 106, App
45	39	34.2	876	7	US-11-077-550-108	Sequence 108, App

ALIGNMENTS

```
RESULT 1
US-11-094-519A-41
; Sequence 41, Application US/11094519A
; Publication No. US20050281810A1
; GENERAL INFORMATION:
; APPLICANT: BERNSTEIN, Jeanne
; APPLICANT: LEVINE, Zurit
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0140P
; CURRENT APPLICATION NUMBER: US/11/094,519A
; CURRENT FILING DATE: 2005-03-31
; PRIOR APPLICATION NUMBER: US/09/695,293
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: IL 132558
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 41
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-094-519A-41

Query Match      43.0%; Score 49; DB 7; Length 865;
Best Local Similarity 45.0%; Pred. No. 4.8;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY      4 PIVAVIGICAVFLQQYV 23
DB      544 PIVAVIGIAVFLVITV 563

RESULT 2
US-10-467-657-4924
; Sequence 4924, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MORACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
```

NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqMan9, version 1.04
SEQ ID NO 4924
LENGTH: 208
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4924

Query Match 39.5%; Score 45; DB 6; Length 208;
Best Local Similarity 61.5%; Pred. No. 4.3;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 PIVAVIGICAAV 16
DB 20 PIALQVLGICSL 32

RESULT 3
US-10-467-657-2722
Sequence 2722, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:

APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqMan9, version 1.04
SEQ ID NO 2722
LENGTH: 759
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2722

Query Match 39.5%; Score 45; DB 6; Length 759;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 PIVAVIGICAAVPL 19
DB 468 PLPLVLAACAALLL 483

RESULT 4
US-11-119-683-3
Sequence 3, Application US/11119683
Publication No. US20050262598A1
GENERAL INFORMATION:

APPLICANT: Gaxiola, Roberto A.
APPLICANT: Fink, Gerald R.
TITLE OF INVENTION: Proton Transporters And Uses In Plants
FILE REFERENCE: 0399.2004-002
CURRENT APPLICATION NUMBER: US/11/119,683
CURRENT FILING DATE: 2005-05-02
PRIOR APPLICATION NUMBER: US/09/834,998
PRIOR FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 09/644,039
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: US 60/164,808
PRIOR FILING DATE: 1999-11-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 633
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Yeast - Schx1
US-11-119-683-3

Query Match 38.6%; Score 44; DB 7; Length 633;
Best Local Similarity 47.8%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 5; Indels 4; Gaps 1;

QY 5 IVAVIGIC---AAVPLQGV 23
DB 379 IIVAAISICVARMCAVPLSQFV 401

RESULT 5
US-10-467-657-842

Sequence 842, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqMan9, version 1.04
SEQ ID NO 842
LENGTH: 98
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-842

Query Match 37.7%; Score 43; DB 6; Length 98;
Best Local Similarity 64.3%; Pred. No. 3.9;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 5 IVAVIGICAAVPL 18
DB 15 IIVAVIGIILAAVNL 28

RESULT 6
US-11-052-554A-238

Sequence 238, Application US/11052554A
Publication No. US20050288866A1
GENERAL INFORMATION:
APPLICANT: Sachdeva, et al.
TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
FILE REFERENCE: 30853/40359A
CURRENT APPLICATION NUMBER: US/11/052,554A
CURRENT FILING DATE: 2005-02-07
PRIOR APPLICATION NUMBER: US 60/589,227
PRIOR FILING DATE: 2004-07-20
PRIOR APPLICATION NUMBER: IN 173/DEL/2004
PRIOR FILING DATE: 2004-02-06
NUMBER OF SEQ ID NOS: 763
SOFTWARE: PatentIn version 3.3
SEQ ID NO 238
LENGTH: 170
TYPE: PRT
ORGANISM: Neisseria meningitidis 22491
US-11-052-554A-238

Query Match 37.7%; Score 43; DB 7; Length 170;
Best Local Similarity 64.3%; Pred. No. 6.9;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;


```

/ TITLE OF INVENTION: genus Tagetes as feedstuffs
/ FILE REFERENCE: 13173-00004-US
/ CURRENT APPLICATION NUMBER: US/10/524,647
/ CURRENT FILING DATE: 2005-02-17
/ PRIOR APPLICATION NUMBER: PCT/EP2003/009109
/ PRIOR FILING DATE: 2003-08-18
/ PRIOR APPLICATION NUMBER: DE 102 38 980.2
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 38 978.0
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 38 979.9
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 53 112.9
/ PRIOR FILING DATE: 2002-11-13
/ PRIOR APPLICATION NUMBER: DE 102 58 971.2
/ PRIOR FILING DATE: 2002-12-16
/ NUMBER OF SEQ ID NOS: 142
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 25
/ LENGTH: 315
/ TYPE: PRT
/ ORGANISM: Haematococcus pluvialis
US-10-524-647-25

Query Match          36.4%; Score 41.5; DB 6; Length 315;
Best Local Similarity 52.2%; Pred. No. 22;
Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 3 PP-----IVLAVIGICAAVFL 18
DB 47 PPSDTKGITMALAVIGSMAAVFL 69

RESULT 14
US-10-524-647-23
/ Sequence 23, Application US/10524647
/ Publication No. US20050281909A1
/ GENERAL INFORMATION:
/ APPLICANT: Flachmann, Ralf
/ APPLICANT: Sauer, Marc
/ APPLICANT: Schopfer, Christel R.
/ APPLICANT: Klebattel, Martin
/ APPLICANT: Pfeiffer, Angelika-Maria
/ APPLICANT: Luck, Thomas
/ APPLICANT: Voeste, Dirk
/ TITLE OF INVENTION: Use of asexanthin-containing plants or parts of plants of the
/ TITLE OF INVENTION: genus Tagetes as feedstuffs
/ FILE REFERENCE: 13173-00004-US
/ CURRENT APPLICATION NUMBER: US/10/524,647
/ CURRENT FILING DATE: 2005-02-17
/ PRIOR APPLICATION NUMBER: PCT/EP2003/009109
/ PRIOR FILING DATE: 2003-08-18
/ PRIOR APPLICATION NUMBER: DE 102 38 980.2
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 38 978.0
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 38 979.9
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 53 112.9
/ PRIOR FILING DATE: 2002-11-13
/ PRIOR APPLICATION NUMBER: DE 102 58 971.2
/ PRIOR FILING DATE: 2002-12-16
/ NUMBER OF SEQ ID NOS: 142
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 23
/ LENGTH: 329
/ TYPE: PRT
/ ORGANISM: Haematococcus pluvialis
US-10-524-647-23

Query Match          36.4%; Score 41.5; DB 6; Length 329;
Best Local Similarity 52.2%; Pred. No. 23;
Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

```

```

QY 3 PP-----IVLAVIGICAAVFL 18
DB 61 PPSDTKGITMALAVIGSMAAVFL 83

RESULT 15
US-10-524-647-80
/ Sequence 80, Application US/10524647
/ Publication No. US20050281909A1
/ GENERAL INFORMATION:
/ APPLICANT: Flachmann, Ralf
/ APPLICANT: Sauer, Marc
/ APPLICANT: Schopfer, Christel R.
/ APPLICANT: Klebattel, Martin
/ APPLICANT: Pfeiffer, Angelika-Maria
/ APPLICANT: Luck, Thomas
/ APPLICANT: Voeste, Dirk
/ TITLE OF INVENTION: Use of asexanthin-containing plants or parts of plants of the
/ TITLE OF INVENTION: genus Tagetes as feedstuffs
/ FILE REFERENCE: 13173-00004-US
/ CURRENT APPLICATION NUMBER: US/10/524,647
/ CURRENT FILING DATE: 2005-02-17
/ PRIOR APPLICATION NUMBER: PCT/EP2003/009109
/ PRIOR FILING DATE: 2003-08-18
/ PRIOR APPLICATION NUMBER: DE 102 38 980.2
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 38 978.0
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 38 979.9
/ PRIOR FILING DATE: 2002-08-20
/ PRIOR APPLICATION NUMBER: DE 102 53 112.9
/ PRIOR FILING DATE: 2002-11-13
/ PRIOR APPLICATION NUMBER: DE 102 58 971.2
/ PRIOR FILING DATE: 2002-12-16
/ NUMBER OF SEQ ID NOS: 142
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 80
/ LENGTH: 329
/ TYPE: PRT
/ ORGANISM: Haematococcus pluvialis
US-10-524-647-80

Query Match          36.4%; Score 41.5; DB 6; Length 329;
Best Local Similarity 52.2%; Pred. No. 23;
Matches 12; Conservative 1; Mismatches 3; Indels 7; Gaps 1;

QY 3 PP-----IVLAVIGICAAVFL 18
DB 61 PPSDTKGITMALAVIGSMAAVFL 83

Search completed: January 23, 2006, 11:53:36
Job time : 5.1 secs

```

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds

(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-8

Sequence: 1 NLPPIVLAVIGICAAVFLQGYV 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*
9: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	23	6	ADBI6893 Rhizobium
2	114	100.0	23	9	AEBO8225 Mesorhizo
3	114	100.0	30	6	ADBI6913 Bacteri
4	114	100.0	30	9	AEBO8247 Bacteri
5	58	50.9	381	5	AAU93052 Arabidops
6	58	50.9	381	7	ADK37211 Plant yie
7	58	50.9	381	8	ADY23705 Plant tra
8	57	50.0	345	8	ADY23705 Plant ful
9	54	47.4	337	8	ADK66009 Plant ful
10	54	47.4	999	8	ADN24218 Bacteri
11	52	45.6	356	3	ADN42360 Arabidops
12	52	45.6	375	5	AAU93173 Arabidops
13	51	44.7	476	7	ADP04975 Bacteri
14	50.5	44.3	746	8	ADK6040 Kluverom
15	50	43.9	246	8	ABO64463 Klebsiell
16	50	43.9	584	6	ABR52244 Rat gluta
17	50	43.9	604	6	ABG74133 Rat metab
18	50	43.9	909	8	ADQ28095 Mouse nov
19	50	43.9	912	7	ADK58164 Rat Prote
20	50	43.9	912	7	ADK58164 Rat Prote
21	49.5	43.4	88	3	AAU93056 Arabidops
22	49.5	43.4	88	3	AAU93245 Arabidops
23	49.5	43.4	88	3	AAU93245 Arabidops
24	49.5	43.4	134	3	AAU93244 Arabidops

25	49.5	43.4	146	3	AAU93055 Arabidops
26	49.5	43.4	146	3	AAU93052 Arabidops
27	49.5	43.4	146	8	ADN74025 Thale cre
28	49.5	43.4	791	8	ABM84773 Human dia
29	49.5	43.4	831	4	AAU93264 Human pol
30	49.5	43.4	831	8	ADQ20117 Human sof
31	49.5	43.4	831	8	ADQ91460 Amino aci
32	49.5	43.4	831	9	ADY70286 Human bet
33	49.5	43.4	831	9	ADY70286 Human pol
34	49.5	43.4	838	4	AAU41050 Human dia
35	49.5	43.4	840	8	ABM84774 Human dia
36	49.5	43.4	880	6	ADL46155 Myline so
37	49	43.0	23	6	ADBI6887 Pasteurel
38	49	43.0	23	9	AEBO8219 Pasteurel
39	49	43.0	25	6	ADBI6898 Kingella
40	49	43.0	25	9	AEBO8230 Kingella
41	49	43.0	203	9	AEBO8262 Pasteurel
42	49	43.0	210	2	AAU72099 Human mgl
43	49	43.0	364	7	ADM05927 Human pro
44	49	43.0	461	8	ADP29713 Human sec
45	49	43.0	481	2	AAU72095 Human mgl

ALIGNMENTS

RESULT 1
ADBI6893
ID ADBI6893 standard; peptide; 23 AA.
XX
AC ADBI6893;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rhizobium loti MCR0187 penetrating peptide 8.
XX
KW penetrating peptide; epithelial; endothelial; tight junction; diabetes;
KW infertility; hormone; vitamin deficiency; neurodegenerative;
KW cardiovascular; haematological; endocrine disorder; obesity;
KW neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
KW osteopathic; cytostatic; nootropic.
XX
OS Mesorhizobium loti.
XX
PN WC02003066859-A2.
XX
PD 14-AUG-2003.
XX
PF 07-FEB-2003; 2003WC-IB000968.
XX
PR 07-FEB-2002; 2002US-0355396P.
XX
PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Ben-Sasson SA, Cohen E;
XX
DR WPI; 2003-697452/66.
XX
PT New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
PS Claim 2; Page 14; 60pp; English.
XX
CC This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 CC cytostatic or neurotropic activities. This peptide is from MLR0187 of
 CC Rhizobium loti and is penetrating peptide 8 of the invention.

XX Sequence 23 AA;
 SQ

Query Match 100.0%; Score 114; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLPPVIAVIGICAAVFLQQYV 23
 Db 1 NLPPVIAVIGICAAVFLQQYV 23

RESULT 2
 AEB08225
 ID AEB08225 standard; peptide; 23 AA.
 AC AEB08225;
 XX 25-AUG-2005 (first entry)
 DT Mesorhizobium loti MLR0187 penetrating peptide 8, SEQ ID NO: 8.
 DE
 XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antipneumatic; cystostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.

XX Mesorhizobium loti.
 OS
 XX US2005136103-A1.
 PN
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 PF
 XX 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SANSSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 XX WPI, 2005-444089/45.
 DR
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 54; SEQ ID NO 8; 59pp; English.

XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Mesorhizobium loti MLR0187
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.

XX Sequence 23 AA;
 SQ

Query Match 100.0%; Score 114; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLPPVIAVIGICAAVFLQQYV 23
 Db 1 NLPPVIAVIGICAAVFLQQYV 23

RESULT 3
 ADB16913
 ID ADB16913 standard; peptide; 30 AA.
 AC ADB16913;
 XX 20-NOV-2003 (first entry)
 DT
 XX Escherichia coli peptide 3 coupled to imaging compound linker, IBW-003.
 DE
 XX epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiant; antiarteriosclerotic; osteopathic; cytostatic; neurotropic;
 KW imaging linker; penetrating peptide; IBW-003.
 XX
 XX Synthetic.
 OS
 XX Escherichia coli.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT Modified-site 30
 FT Modified-site 30
 FT /note= "C-terminal amide"
 XX WO2003066859-A2.
 PN
 XX 14-AUG-2003.
 PD
 XX 07-FEB-2003; 2003WO-IB000968.
 PF
 XX 07-FEB-2002; 2002US-0355396P.
 PR
 XX

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen B;
 PI WPI, 2003-697452/66.
 XX
 XX
 PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 XX Example 3, Page 40; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cytosolic or neurotropic activities. This peptide sequence is IBM-003,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 XX
 SO Sequence 30 AA:
 Query Match 100.0%; Score 114; DB 6; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.8e-10; Indels 0; Gaps 0;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NLPPIYLAIVGICAVFLQQYV 23
 1 NLPPIYLAIVGICAVFLQQYV 23
 DB
 RESULT 4
 AEB08247
 ID AEB08247 standard; peptide; 30 AA.
 XX
 XX AEB08247;
 XX
 DT 25-AUG-2005 (first entry)
 DE Penetrating peptide SEQ: 30 used in composition for mucosal vaccination.
 XX
 XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX anti-diabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; ophthalmologic;
 XX degeneration; musculoskeletal disease; ocular disease;
 XX neurodegenerative disease; neuroprotective; alzheimers disease;
 XX neurotropic; neurological disease; parkinsons disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotoxic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency; infection;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; vinticide;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antirheumatic; cytosolic;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.

XX
 OS Unidentified.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT Misc-difference 1
 FT /note= "N-terminal acylated"
 FT Misc-difference 30
 FT /note= "Optionally C-terminal amide, optionally the free
 FT amino group of lysine is acylated with a fatty acid"
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHR/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen B;
 XX WPI, 2005-444089/45.
 XX
 XX
 XX Claim 63; SEQ ID NO 30; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, autoimmune diseases, immune deficiencies,
 CC rheumatologic disorders, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is a penetrating peptide (IBM-003) used
 CC in the composition for mucosal vaccination using a counter anion and a
 CC penetrating peptide.
 XX
 XX Sequence 30 AA:
 Query Match 100.0%; Score 114; DB 9; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.8e-10; Indels 0; Gaps 0;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NLPPIYLAIVGICAVFLQQYV 23
 1 NLPPIYLAIVGICAVFLQQYV 23
 DB
 RESULT 5
 AAU93052
 ID AAU93052 standard; protein; 381 AA.
 XX
 XX AAU93052;

02-JUL-2002 (first entry)

Arabidopsis transcription factor #90.

Agriculture; metabolic chemical; environmental stress; drought; microbial disease resistance; herbicide resistance; seed yield; fruit yield; growth rate; leaf senescence; flower senescence. plant; transcription factor; transgenic.

Arabidopsis thaliana.

WO200215675-A1.

28-FEB-2002.

22-AUG-2001; 2001WO-US026189.

22-AUG-2000; 2000US-0227439P.
16-NOV-2000; 2000US-00713994.
18-APR-2001; 2001US-00837944.

(MEND-) MENDEL BIOTECHNOLOGY INC.
(PILG/) PILGRIM M.
(CREB/) CREBLMAN R.
(DUBE/) DUBELT A J.
(HEAR/) HEARD J.
(JIANG/) JIANG C.
(KEDD/) KEDDIE J.
(ADAM/) ADAM L.
(RATC/) RATCIEP O.
(REUB/) REUBER J L.
(RIEC/) RIECHMANN J L.
(YUGG/) YU G.
(PINED/) PINEDA O.

Pilgrim M, Creelman R, Dubell AJ, Heard J, Jiang C, Keddie J, Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O; N-PSDB, ABK65238.

WPI; 2002-292022/33.

An isolated or recombinant polynucleotide used to produce a transgenic plant.

Claim 40; Page 398-400; 941pp; English.

The invention relates to 1 of 232 isolated or recombinant polynucleotides encoding an Arabidopsis thaliana transcription factor, their variants, complements, fragments, or related polynucleotide with 31% to 95% sequence identity, where the plant possesses an altered trait as compared to a wild-type or reference plant, or the plant exhibits an altered phenotype as compared to a wild-type or reference plant, or the plant exhibits ectopic expression or altered expression of one or more genes associated with a plant trait as compared to a wild plant. Also included are a transgenic plant comprising the polynucleotides, a computer readable medium having stored sequence information, and identifying a homologue sequence from a database comprising a plurality of known plant sequences comprising inputting sequence information selected from one of 464 fully defined sequences given in the specification. The isolated or recombinant polynucleotide is used for producing a plant having a modified trait, the method comprising selecting a polynucleotide that encodes a polypeptide or an antisense nucleic acid, inserting the polynucleotide or antisense nucleic acid into an expression vector, introducing the vector into a plant or a cell of a plant to overexpress the polypeptide or antisense nucleic acid, thereby producing a modified plant, and selecting for a modified trait (e.g. increased production of agriculturally useful proteins or metabolic chemical, pest tolerance, environmental stress response (e.g. drought), microbial disease resistance, herbicide resistance, seed and fruit yield, growth rate, leaf senescence and many other traits listed in the specification). The present invention is one of the 232 proteins which are A. thaliana transcription factors

Query Match	50.9%;	Score 58;	DB 5;	Length 381;
Best Local Similarity	54.5%;	Pred. No. 2.1;		
Matches 12;	Conservative	2;	Mismatches 8;	Indels 0; Gaps 0
Qy	1	NLPPIVLAVIGCAVFLPQQY 22		
Db	40	NFPIAIVIGLAVPFLVSY 61		
RESULT 6				
ADBE37211				
ID	ADBE37211	standard;	protein;	381 AA.
AC	ADBE37211;			
XX				
DT	29-JAN-2004	(first entry)		
XX				
DE	Plant yield related protein from clone GI095.			
XX				
KM	transcription factor; tolerance; environmental condition;			
KM	microbial disease; fungal disease; viral disease; pest infestation;			
KM	herbicide sensitivity; heavy metal tolerance; heavy metal uptake;			
KM	growth improvement; photocondition; nutrient uptake; hormone sensitivity;			
XX	transgenic plant.			
XX				
OS	Arabidopsis thaliana.			
XX				
PN	MO2003014327-A2.			
PD	20-FEB-2003.			
XX				
PF	09-AUG-2002; 2002MO-US026966.			
XX				
PR	09-AUG-2001; 2001US-0310847P.			
PR	19-NOV-2001; 2001US-0336049P.			
PR	11-DEC-2001; 2001US-0336692P.			
PR	14-JUN-2002; 2002US-00171468.			
XX				
PA	(MEND-) MENDEL BIOTECHNOLOGY INC.			
PI	Reuber TL, Riechmann JT, Heard JE, Jiang C, Adam LJ, Dubell AN,			
PI	Ratcliffe O, Pineda O, Yu GL, Brown PG,			
XX				
DR	WPI: 2003-256576/25.			
DR	N-PSDB; ADB37210.			
XX				
PT	New stress-related transcription factor polynucleotides and polypeptides,			
PT	useful for producing transgenic plants with e.g. improved tolerance to			
PT	diseases or pests, decreased herbicide sensitivity, or improved nutrient			
PT	uptake.			
XX				
PS	Disclosure; SEQ ID NO 150; 470bp; English.			
XX				
CC	The invention relates to a number of cDNA sequence and their encoded			
CC	proteins which are especially transcription factor cDNAs and their			
CC	proteins. The isolated or recombinant polynucleotide is useful for			
CC	producing a modified plant with a modified trait, e.g. enhanced tolerance			
CC	to environmental conditions; improved tolerance to microbial, fungal or			
CC	viral diseases; improved tolerance to pest infestation; decreased			
CC	herbicide sensitivity; improved tolerance of heavy metals; or enhanced			
CC	ability to take up heavy metals; improved growth under poor			
CC	photoconditions; improved nutrient uptake; or reduced hormone			
CC	sensitivity. The transgenic plants are useful for growing a progeny plant			
CC	comprising the desired trait. The polynucleotides and polypeptides are			
CC	also useful in bioinformatic search methods. This sequence represents one			
XX	of the proteins of the invention.			
XX				
SO	Sequence 381 AA;			

CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

CC Sequence 345 AA;

CC Query Match 50.0%; Score 57; DB 8; Length 345;

CC Best Local Similarity 52.6%; Pred. No. 2.7;

CC Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

CC 4 PIVLAVIGICAAVFLQQY 22

DB 52 PVLVILIGLMSAFLLVSY 70

RESULT 9

ID ADX66009 standard; protein; 327 AA.

AC ADX66009;

DT 21-APR-2005 (first entry)

DE Plant full length insert polypeptide seqid 36852.

XX plant protectant; plant growth regulator; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.

OS Unidentified.

PN US2004034888-A1.

PD 19-FEB-2004.

PF 28-APR-2003; 2003US-00425114.

PR 06-MAY-1999; 98US-00304517.

PR 05-NOV-2001; 2001US-00985678.

XX (LIUJ/) LIU J.

PA (ZHOU/) ZHOU Y.

PA (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S E.

PA (TABAS/) TABASKA J E.

PA (CAOY/) CAO Y.

XX LIU J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

DR WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to

PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or

PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.

PS Claim 1; SEQ ID NO 36852; 15pp; English.

CC The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp://seqdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

CC Sequence 327 AA;

CC Query Match 47.4%; Score 54; DB 8; Length 327;

CC Best Local Similarity 50.0%; Pred. No. 7.5;

CC Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

CC 1 NLPPVILAVIGICAAVFLQQY 22

DB 3 SFPVILAVIGICAAVFLVSY 24

RESULT 10

ID ADN24218 standard; protein; 999 AA.

AC ADN24218;

DT 02-DEC-2004 (first entry)

DE Bacterial polypeptide #6871.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

OS Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX CAO Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

DR WPI; 2004-061375/06.

XX New recombinant DNA construct, useful for improving plant tolerance to

PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or

PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-01443352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151388P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154799P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158932P.
PR 12-OCT-1999; 99US-0159369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.

PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161931P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 45.6%; Score 52; DB 3; Length 356;
Best Local Similarity 47.4%; Pred. No. 17;
Matches 9; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 4 PIVLAIVIGICAAVPLLOOY 22
|:::|::|:|:|
Db 58 PLRLALIGILASAFILSVY 76

RESULT 12
AAU93173
ID AAU93173 standard; protein; 375 AA.
AC AAU93173;
XX 02-JUL-2002 (first entry)
XX Arabidopsis transcripion factor #211.
DE Arabidopsis transcripion factor #211.
XX
XX Agriculture; metabolic chemical; environmental stress; drought;
XX microbial diseasee resistance; herbicide resistance; seed yield;
XX fruit yield; growth rate; leaf senescence; flower senescence. plant;
XX transcripion factor; transgenic.
OS Arabidopsis thaliana.
XX
XX WO200215675-A1.
XX
XX 28-FEB-2002.
XX
XX 22-AUG-2001; 2001WO-US026189.
XX
XX 22-AUG-2000; 2000US-0227439P.
XX 16-NOV-2000; 2000US-00713994.
XX 18-APR-2001; 2001US-00837944.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX (PIUG/) PILGERIM M.
XX (CREE/) CREELMAN R.
XX (DUBE/) DOBELLE A J.
XX (HEAR/) HEARD J.
XX (JIANG/) JIANG C.
XX (KEDD/) KEDDIE J.
XX (ADAM/) ADAM L.
XX (RATC/) RATCLIFF O.
XX (REUB/) REUBER J L.
XX (RIEC/) RIECHMANN J L.

PA (YUG3/) YU G.
 PA (PINE/) PINEDA O.
 PI Pilgrim M., Creelman R., Dubell AJ, Heard J, Jiang C, Keddie J,
 PI Adam L, Ratcliff O, Reuber JL, Riechmann JL, Yu G, Pineda O;
 XX WPI; 2002-292022/33.
 DR N-PSDB; ABR65359.
 XX
 PT An isolated or recombinant polynucleotide used to produce a transgenic
 PT plant.
 XX
 PS Claim 40; Page 868-870; 941pp; English.
 XX
 CC The invention relates to 1 of 232 isolated or recombinant polynucleotides
 CC encoding an Arabidopsis thaliana transcription factor, their variants,
 CC complements, fragments, or related polynucleotide with 31% to 95%
 CC sequence identity, where the plant possesses an altered trait as compared
 CC to a wild-type or reference plant, or the plant exhibits an altered
 CC phenotype as compared to a wild-type or reference plant, or the plant
 CC exhibits ectopic expression or altered expression of one or more genes
 CC associated with a plant trait as compared to a wild plant. Also included
 CC are a transgenic plant comprising the polynucleotides, a computer
 CC readable medium having stored sequence information, and identifying a
 CC homologous sequence from a database comprising a plurality of known plant
 CC sequences comprising inputting sequence information selected from one of
 CC 464 fully defined sequences given in the specification. The isolated or
 CC recombinant polynucleotide is used for producing a plant having a
 CC modified trait, the method comprising selecting a polynucleotide that
 CC encodes a polypeptide or an antisense nucleic acid, inserting the
 CC polynucleotide or antisense nucleic acid into an expression vector,
 CC introducing the vector into a plant or a cell of a plant to overexpress
 CC the polypeptide or antisense nucleic acid, thereby producing a modified
 CC plant, and selecting for a modified trait (e.g. increased production of
 CC agriculturally useful proteins or metabolic chemicals, pest tolerance,
 CC environmental stress response (e.g. drought), microbial disease
 CC resistance, herbicide resistance, seed and fruit yield, growth rate, leaf
 CC and flower senescence and many other traits listed in the specification).
 CC The present sequence is one of the 232 proteins which are A. thaliana
 CC transcription factors
 CC
 XX
 SQ Sequence 375 AA;
 XX
 Query Match 45.6%; Score 52; DB 5; Length 375;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MLRPYLVAVIGICAAVFLQOY 22
 Db 33 SPPIVAVAVIGILATPLAVSY 54
 XX
 RESULT 13
 ID ADF04975
 XX ADF04975 standard; protein; 476 AA.
 XX
 AC ADF04975;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 XX Bacterial polypeptide #1088.
 DE
 XX
 KM Proteus mirabilis infection; bacterial infection; antibacterial;
 KM immunostimulant.
 XX
 KM Proteus mirabilis.
 OS
 XX
 XX Proteus mirabilis.
 OS
 XX
 PN US6605709-B1.
 XX
 PD 12-AUG-2003.
 XX
 PP 05-APR-2000; 2000US-00543681.
 XX

PR 09-APR-1999; 99US-0128706P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX Breton GL;
 PI
 XX WPI; 2003-895291/82.
 DR N-PSDB; ADF00803.
 XX
 PT New Proteus mirabilis polypeptides and polynucleotides, useful as
 PT reagents for diagnosis of bacterial disease, as components of
 PT antibacterial vaccines, as targets for antibacterial drugs, or as
 PT biocontrol agents for plants.
 XX
 PS Disclosure; SEQ ID NO 5260; 870pp; English.
 XX
 CC The invention relates to new Proteus mirabilis polypeptides and
 CC polynucleotides. The invention also relates to antibodies against the
 CC polypeptides, methods for producing the polypeptides, a method of
 CC generating vaccines for immunizing an individual against P. mirabilis, a
 CC method for evaluating a compound for the ability to bind a P. mirabilis
 CC polypeptide and a method for screening test compounds for anti-bacterial
 CC activity. The polypeptides and polynucleotides are useful as molecular
 CC targets for diagnosing, preventing and treating pathological conditions
 CC resulting from bacterial infection, as reagents for diagnosis of
 CC bacterial diseases, as components of antibacterial vaccines, as targets
 CC for antibacterial drugs or as bio-control agents for plants. This
 CC sequence represents a Proteus mirabilis polypeptide of the invention.
 CC
 XX
 SQ Sequence 476 AA;
 XX
 Query Match 44.7%; Score 51; DB 7; Length 476;
 Best Local Similarity 40.9%; Pred. No. 33;
 Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
 QY 2 LPPIYLVAVIGICAAVFLQOY 23
 Db 221 LPISIAIIVGMCIHFFMORYL 242
 XX
 RESULT 14
 ID ADR46040
 XX ADR46040 standard; protein; 746 AA.
 XX
 AC ADR46040;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Kluyveromyces lactis MNM1 protein SEQ ID NO:44.
 XX
 KM Kluyveromyces lactis MNM1 protein SEQ ID NO:44.
 KM human-like glycoprotein; glycoprotein; N-glycan;
 KM MANS5G1CNC2 carbohydrate structure; MANS5G1CNC2; glycosylation enzyme;
 KM cellular targeting signal peptide; catalytic domain region; mannosidase;
 KM glycosyltransferase; glycosidase; MNM1.
 XX
 OS Kluyveromyces lactis.
 XX
 XX WO2004074499-A2.
 PN
 XX
 PD 02-SEP-2004.
 XX
 XX 20-FEB-2004; 2004MO-US005244.
 PP
 XX 20-FEB-2003; 2003US-00371877.
 PR
 XX (GERN/) GERNGROSS T U.
 PA (WILD/) WILDT S.
 PA (CHOI/) CHOI B.
 PA (NETT/) NETT J H.
 PA (BOBR/) BOBROWICZ P.
 PA (HAML/) HAMILTON S R.
 PA (DAVI/) DAVIDSON R C.
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 Seconds
(without alignments)
267.334 Million cell updates/sec

Title: US-10-501-838a-8
Sequence: 1 NLPPIVAVIGICAAVFLQQYV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	237	Q98ND6_RHILO	Q98nd6 rhizobium 1
2	74	64.9	235	Q8YHQ7_BRUMB	Q8yhg7 bruceella me
3	74	64.9	253	Q57CN7_BRUBB	Q57cn7 bruceella ab
4	74	64.9	253	Q8G059_BRUBU	Q8g059 bruceella su
5	62	54.4	336	ATL4A_ARATH	ATL4a arabidopsis
6	58	50.9	381	ATL1A_ARATH	ATL1a arabidopsis
7	58	50.9	383	Q7XLY8_ORYSA	Q7xly8 oryza sativ
8	58	50.9	389	Q6Z8A6_ORYSA	Q6z8a6 oryza sativ
9	58	50.9	389	Q6Z225_CABBR	Q6z225 caenorhabdi
10	54	47.4	162	Q6Z151_ORYSA	Q6z151 oryza sativ
11	54	47.4	999	MGR1_CAREL	Q09630 caenorhabdi
12	53	46.5	278	Q5WCT6_BACSK	Q5wct6 bacillus cl
13	53	46.5	285	Q6AHM4_LEIKX	Q6ahm4 leifsonia x
14	53	46.5	429	Q6SN76_BACLD	Q6sn76 bacillus 11
15	52	45.6	107	Q5YB9_NOCFA	Q5yb9 nocardia fa
16	52	45.6	157	Q7O5Z8_ANOGA	Q7o5z8 anopheles g
17	52	45.6	356	ATL3A_ARATH	ATL3a arabidopsis
18	52	45.6	375	ATL5K_ARATH	ATL5k arabidopsis
19	51	44.7	134	Q5FR38_GLUOX	Q5fr38 gluconobact
20	51	44.7	256	Q7UIH5_RHOBA	Q7uih5 rhodospirill
21	51	44.7	311	Q8H077_ORYSA	Q8h077 oryza sativ
22	51	44.7	435	Q6MCP8_PARIW	Q6mcp8 parachlamyd
23	51	44.7	519	Q33968_STRCT	Q33968 streptomyce
24	51	44.7	581	Q83AV7_COXBU	Q83av7 coxilla bu
25	51	44.7	953	YMG8_YEAST	Q03516 saccharomyc
26	50.5	44.3	746	Q6CNT7_KIULU	Q6cnt7 kiuyveromyc
27	50	43.9	115	Q80UC2_MOUSE	Q80uc2 mus musculu
28	50	43.9	120	V041_URBPA	Q09ra4 ureaplasma
29	50	43.9	394	Q4QH50_LEIMA	Q4qhg8 leishmania
30	50	43.9	682	Q4S5Q2_TETNG	Q4s5q2 tetradon n
31	50	43.9	715	Q8DGW9_SYNEL	Q8dgv9 bynechococc

32	50	43.9	811	2	Q4RJZ9_TETNG	Q4rjz9 tetradon n
33	50	43.9	832	2	Q6BFA4_MOUSE	Q6bfa4 mus musculu
34	50	43.9	912	1	MGR1_PAT	P31423 rattus norv
35	50	43.9	983	2	Q6Z916_PAT	Q6z916 rattus norv
36	49.5	43.4	134	2	Q49486_ARATH	Q49486 arabidopsis
37	49.5	43.4	146	2	Q9W719_ARATH	Q9w719 arabidopsis
38	49.5	43.4	661	2	Q8K043_MOUSE	Q8k043 mus musculu
39	49.5	43.4	825	2	Q9QXW6_MOUSE	Q9qxw6 mus musculu
40	49.5	43.4	825	2	Q6PHUS_MOUSE	Q6phus mus musculu
41	49.5	43.4	831	1	SORT_HUMAN	Q99523 homo sapien
42	49.5	43.4	831	2	Q81Z49_HUMAN	Q81z49 homo sapien
43	49	43.0	156	2	P77876_YNEIS	P77876 kingella de
44	49	43.0	203	1	Y1850_PASMU	Q9c1y8 pasteurella
45	49	43.0	228	2	Q87PM6_VIBPA	Q87pm8 vibrio para

ALIGNMENTS

RESULT 1	Q98ND6_RHILO	PRT;	237 AA.
AC	Q98ND6;		
DT	01-OCT-2001 (TrEMBLrel. 18, Created)		
DT	01-OCT-2001 (TrEMBLrel. 18, Last sequence update)		
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)		
DE	M1r0187 protein.		
GN	OrderedLocustNames=m1r0187;		
OS	Rhizobium loti (Mesorhizobium loti).		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Phyllobacteriaceae; Mesorhizobium.		
NCBI_TaxId=381;			
RN	[1]		
RP	NCLEOTIDE SEQUENCE.		
RC	STRAIN=MAF30309;		
RX	MEDLINE=21082930; PubMed=11214968;		
RA	Kaneho T., Nakamura Y., Sato S., Aamizu B., Kato T., Sasamoto S.,		
RA	Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,		
RA	Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,		
RA	Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,		
RA	Takeuchi C., Yamada M., Tabata S.		
RT	"Complete genome structure of the nitrogen-fixing symbiotic bacterium		
RT	Mesorhizobium loti."		
RL	DNA_Res-74331-336(2000).		
DR	EMBL; BA000012; BAB47825.1; -, Genomic DNA.		
DR	Interpro; IPR002610; Rhomboid_1ike.		
DR	Pfam; PF01694; Rhomboid; 1.		
KW	Complete proteome.		
SQ	SEQUENCE 237 AA; 25745 MW; E3BCE8PDC5C6C53 CRC64;		
Query Match	100.0%; Score 114; DB 2; Length 237;		
Best Local Similarity	100.0%; Pred. No. 4,2e+08;		
Matches	23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 NLPPIVAVIGICAAVFLQQYV 23		
DB	11 NLPPIVAVIGICAAVFLQQYV 33		
RESULT 2	Q8YHQ7_BRUMB	PRT;	235 AA.
ID	Q8YHQ7_BRUMB PRELIMINARY;		
AC	Q8YHQ7;		
DT	01-MAR-2002 (TrEMBLrel. 20, Created)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	INTRACELLULAR MEMBRANE PROTEIN (Rhomboid family).		
GN	OrderedLocustNames=BME10739;		
OS	Brucella melitensis.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Brucellaceae; Brucella.		
NCBI_TaxId=29459;			
RN	[1]		

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX PubMed=11756688; DOI=10.1073/pnas.221575398;
RA Delvecchio V.G., Kapetral V., Redkar R.J., Patra G., Muijer C., Los T.,
RA Ivanova N., Anderson N., Bhattacharyya A., Lykdis A., Reznik G.,
RA Jablonki L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elser P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haseklov R., Kyprides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009515; AAL51920.1; -; Genomic_DNA.
DR PIR; AE3344; AE3344.
DR InterPro; IPR002610; Rhomboid_1like.
DR Pfam; PF01694; Rhomboid; 1.
KM Complete proteome.
SQ SEQUENCE 235 AA; 25522 MW; CFE0187A99DC1B0 CRC64;

Query Match 64.9%; Score 74; DB 2; Length 235;
Best Local Similarity 47.8%; Pred. No. 0.022;
Matches 11; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NLPPIVAVIGICAAVFLQGY 23
Db 12 NIPGVVIALIGCAVAVYQNYI 34

RESULT 3
ID Q57CN7_BRUBA PRELIMINARY; PRT; 253 AA.
AC Q57CN7;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Rhomboid family protein.
GN OrderedLocNames=BrubA1_1259;
OS Brucella abortus.
OC Bacteri; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=9-941 / Biovar 1;
RX PubMed=15805518; DOI=10.1128/JB.187.8.2715-2726.2005;
RA Halling S.M., Peterson-Burch B.D., Bricker B.J., Zuercher R.L.,
RA Qing Z., Li L.-L., Kapur V., Alt D.P., Olsen S.C.;
RT "Completion of the genome sequence of Brucella melitensis and comparison
RT to the highly similar genomes of Brucella melitensis and Brucella
RT suis.";
RL J. Bacteriol. 187:2715-2726(2005).
DR EMBL; AE017223; AAX74597.1; -; Genomic_DNA.
KM Complete proteome.
SQ SEQUENCE 253 AA; 27715 MW; A9D915ACBC133B1C CRC64;

Query Match 64.9%; Score 74; DB 2; Length 253;
Best Local Similarity 47.8%; Pred. No. 0.024;
Matches 11; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NLPPIVAVIGICAAVFLQGY 23
Db 30 NIPGVVIALIGCAVAVYQNYI 52

RESULT 4
ID Q8G059_BRUSU PRELIMINARY; PRT; 253 AA.
AC Q8G059;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 23, Last annotation update)
DE Rhomboid family protein.
GN OrderedLocNames=BR1256;
OS Brucella suis.

```

```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tetteijn H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014291; AAN30174.1; -; Genomic_DNA.
DR TIGR; BR1256; -.
DR InterPro; IPR002610; Rhomboid_1like.
DR Pfam; PF01694; Rhomboid; 1.
KM Complete proteome.
SQ SEQUENCE 253 AA; 27784 MW; A8720F1CEC133B1C CRC64;

Query Match 64.9%; Score 74; DB 2; Length 253;
Best Local Similarity 47.8%; Pred. No. 0.024;
Matches 11; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NLPPIVAVIGICAAVFLQGY 23
Db 30 NIPGVVIALIGCAVAVYQNYI 52

RESULT 5
ID ATL4H_ARATH STANDARD; PRT; 336 AA.
AC POC041; Q49690;
DT 13-SEP-2005 (Ref. 48, Created)
DT 13-SEP-2005 (Ref. 48, Last sequence update)
DT 13-SEP-2005 (Ref. 48, Last annotation update)
DE Putative RING-H2 finger protein ATL4H.
GN Name=ATL4H; OrderedLocNames=At4g17910; ORFNames=T6K21.90;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Rosidae; eucoside II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198; DOI=10.1038/47134;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Eutling K.-D., Terry N.,
RA Harris B., Ansoergo W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermayer B., Mache R., Mueller M.,
RA Kreis M., Delzeny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bannroft I.,
RA Vos P., Honelsel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymoprez B., Chung Y.-J., Vandenbusche F.,
RA Braeken M., Weltjens I., Voet M., Bastiens I., Aert R., Defoor E.,
RA Weltzenegger T., Boche G., Rampsperger U., Hilbert H., Brun M.,
RA Holzner E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Moolman P., Klein Lankhorst R., Roze M., Hauf J., Koeter P.,
RA Bernesler S., Hempel S., Feldpausch M., Lambert S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gieles J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA Petter A., Rajandream M.A., Lyne M., Benes V., Reckmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maere A.C., Schaefer M., Meiller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandenath K., Dauner D., Herzl A.,
RA Neumann S., Argitirov A., Vitale D., Ligouri R., Piravandi E.,

```


RA Massenet O., Quigley F., Clabaud G., Muehlen A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chefor F., Cooke R., Berger C., Monfort A., Casachubeta B.,
 RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 RA Perez-Perez A., Parnolle B., Bent B., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bietke C.,
 RA Frichman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Devan M., Wilson R.K., de la Baetle M., Habermann K.,
 RA Parnell L., Dehnan N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mink P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
 RA Nelson J., Spilath J., Ryan B., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonolu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Freeson R., Vill D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Grant S., Shohdy N., Haegawa A., Hameed A., Lodi M., Johnson A.,
 RA Chen B., Marra M.A., Martensen R., McCombie W.R.,
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:769-777(1999).
 RN
 RP GENE FAMILY ORGANIZATION.
 RX MEDLINE=21979369; PubMed=11983057;
 RA Kosarev P., Meyer K.F.X., Hardtke C.S.;
 RT "Evaluation and classification of RING-finger domains encoded by the
 RL Arabidopsis genome.";
 RT Genome Biol. 3:RESEARCH0016.12(2002).
 CC -1- DOMAIN: Ubiquitin conjugation; third step.
 CC -1- PATHWAY: The RING-type zinc finger domain mediates binding to an E2
 CC ubiquitin-conjugating enzyme (By similarity).
 CC -1- SIMILARITY: Belongs to the RING-type zinc finger family. ATL
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -1- CAUTION: Ref.1 sequences differ from that shown due to erroneous
 CC gene model prediction. Was originally assigned as one gene, which
 CC has been split in two genes Atg17910 and Atg17915.
 CC -1- CAUTION: This is a conceptual translation.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 DR EMBL, AL021889; CAA17134.1; ALT SEQ; Genomic DNA.
 DR EMBL, AL161547; CAB78793.1; ALT SEQ; Genomic DNA.
 DR PIR, T05077; T05077.
 DR HSP, Q9LRB7; 11YM.
 DR InterPro: IPR001841; Znf_RING.
 DR Pfam: PF00097; zf-C3HC4; 1.
 DR SMART, SM00184; RING; 1.
 DR PROSITE: PS00518; ZF_RING_2; 1.
 DR PROSITE: PS50089; ZF_RING_2; 1.
 KW Hypothetical protein; metal-binding; Transmembrane;
 KW Ub1 conjugation pathway; Zinc; Zinc-finger.
 FT TRANSMEM 88 108 potential.
 FT ZN FING 181 223 RING-type; atypical.
 SQ SEQUENCE 336 AA; 37415 MW; 2AD68708AF3E08C3 CRC64;

Query Match 54.4%; Score 62; DB 1; Length 336;
 Best Local Similarity 54.5%; Pred. No. 1.6;
 Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

1 NLPYVLAIVIGICAAVFLQOY 22
 83 NLSPLVLAIRGIRFATPLAAV 104

RESULT 6

AT1LA ARATH
 ID AT1LA_ARATH STANDARD; PRT; 381 AA.
 AC P93823;
 DT 13-SEP-2005 (rel. 48; Created)
 DT 13-SEP-2005 (rel. 48; Last sequence update)
 DT 13-SEP-2005 (rel. 48; Last annotation update)
 DE RING-H2 finger protein AT1LA.
 GN Name:AT1LA; OrderedLocustNames=At1G04360; ORFNames=P19P19.21;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_Taxid=3702;
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712; DOI=10.1038/35048500;
 RA Theologis A., Becker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Alcafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.P., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Feng B., Fujii C.Y.,
 RA Gill J.B., Goldsmith A.D., Haas B., Hansen N.P., Hughes B., Hultzer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Klm C.J., Koo H.L., Kremenskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Ienz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malt R., Matzall A.,
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Ritzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tumbunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vayberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Cantini P.,
 RA Chao Q., Choy N., Enji A., Goldsmith A.D., Gurjel M., Hansen N.P.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsu V.W., Iida K., Karnes M.,
 RA Kian S., Koeseema B., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Becker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RL Science 302:842-846(2003).
 RN
 RP GENE FAMILY ORGANIZATION.
 RX PubMed=11983057;
 RA Kosarev P., Meyer K.F.X., Hardtke C.S.;
 RT "Evaluation and classification of RING-finger domains encoded by the
 RL Arabidopsis genome.";
 RT Genome Biol. 3:RESEARCH0016.12(2002).
 CC -1- DOMAIN: Ubiquitin conjugation; third step.
 CC -1- PATHWAY: The RING-type zinc finger domain mediates binding to an E2
 CC ubiquitin-conjugating enzyme (By similarity).
 CC -1- SIMILARITY: Belongs to the RING-type zinc finger family. ATL
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not

```

CC removed.
-----
DR EMBL: AC000104; AAB70441.1; -; Genomic_DNA.
DR EMBL: BT004287; AA042287.1; -; mRNA.
DR EMBL: BT006138; AAP04123.1; -; mRNA.
DR PIR: B86175; B86175.
DR HSSP: Q9LRB7; 11YM.
DR InterPro: IPR001841; Znf_RING.
DR Pfam: PF00097; zf-C3HC4_1.
DR SMART: SM00184; RING_1.
DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE: PS50089; ZF_RING_2; 1.
KW Metal-binding; Transmembrane; Ubl conjugation pathway; Zinc;
KM Zinc-finger.
FT TRANSMEM 46 66 Potential.
FT ZN FING 134 176 RING-type; atypical.
FT COMBINS 16 22 Poly-Pro.
SQ SEQUENCE 381 AA; 42616 MW; 6FA842915F3E796 CRC64;

Query Match
Best Local Similarity 50.9%; Score 58; DB 1; Length 381;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 NPIVLAIVIGICAAVFLIQY 22
Db 40 NPIVLAIVIGILATAFLIVSY 61

RESULT 7
Q7XLY8_ORYSA PRELIMINARY; PRT; 383 AA.
AC Q7XLY8;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DS OSUBA0086006.21 Protein.
GN Name=OSUBA0086006.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE:2237377; PubMed=12447439; DOI=10.1038/nature01183;
RX Pengu O., Zhang Y., Hao P., Wang S., Pu G., Huang Y., Li Y., Zhu J.,
Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
Weng Q., Zhang L., Lu Y., Wu Y., Lu Y., Zhang L.S., Yu Z., Fan D.,
Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
Lan L., Lai Y., Cheng Z., Gu W., Jiang J., Li J., Hong G., Xue Y.,
Han B.;
RT "Sequence and analysis of rice chromosome 4."
RL Nature 420:316-320(2002).
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL: AL662981; CAE04873.2; -; Genomic_DNA.
DR HSSP: Q9LRB7; 11YM.
DR Gramene: Q7XLY8; -.
DR GO: GO:0000151; Cubiquitin ligase complex; IEA.
DR GO: GO:0046872; F:metal ion binding; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0016567; P:protein ubiquitination; IEA.
DR InterPro: IPR001841; Znf_RING.
DR Pfam: PF00097; zf-C3HC4_1.
DR SMART: SM00184; RING_1.
DR PROSITE: PS50089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 383 AA; 40680 MW; F663D2F64063C9FA CRC64;

```

```

Query Match
Best Local Similarity 50.9%; Score 58; DB 2; Length 383;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 4 PIVLAIVIGICAAVFLIQY 22
Db 57 PIVLAIVIGILASAFLLVSY 75

RESULT 8
Q6Z8A6_ORYSA PRELIMINARY; PRT; 389 AA.
AC Q6Z8A6;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DS Zinc finger (C3HC4-type RING finger)-like.
GN Name=P017H1.3; Synonyms=OJ1148_D05.15;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN NUCLEOTIDE SEQUENCE.
RP Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nippobare (GA3) genomic DNA, chromosome 2, PAC
clone: P017H1.3."
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RL NUCLEOTIDE SEQUENCE.
RP Sasaki T., Matsumoto T., Yamamoto K.;
RA "Oryza sativa nippobare (GA3) genomic DNA, chromosome 2, BAC
clone: OJ1148_D05."
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC EMBL: AP004786; BAD07893.1; -; Genomic_DNA.
DR EMBL: AP004118; BAD07693.1; -; Genomic_DNA.
DR Gramene: Q6Z8A6; -.
DR GO: GO:0000151; Cubiquitin ligase complex; IEA.
DR GO: GO:0046872; F:metal ion binding; IEA.
DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0016567; P:protein ubiquitination; IEA.
DR InterPro: IPR001841; Znf_RING.
DR Pfam: PF00097; zf-C3HC4_1.
DR SMART: SM00184; RING_1.
DR PROSITE: PS50089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 389 AA; 41609 MW; A2F331B65F43582E CRC64;

Query Match
Best Local Similarity 50.9%; Score 58; DB 2; Length 389;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 4 PIVLAIVIGICAAVFLIQY 22
Db 70 PIVLAIVIGILASAFLLVSY 88

RESULT 9
Q6Z2Z5_CABR PRELIMINARY; PRT; 963 AA.
AC Q6Z2Z5;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DS Hypothetical protein CBG01949 (Fragment).
GN Name=CBG01949;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.

```

OX NCBI_TaxID=6238;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RG NIH - Mammalian Gene Collection (MGC) project;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL; CAAC0100010; CAB58757.1; -!- Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0008620; C:membrane; IEA.
 DR GO; GO:0008667; P:metabotropic glutamate, GABA-B-like receptor. . . ; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR001828; P:G-protein coupled receptor protein signalin. . . ; IEA.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR InterPro; IPR000162; MrglR_receptor.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR Pfam; PF07562; NCD3G; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PRINTS; PR00593; MTABOTROPIC.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS50259; G_PROTEIN_RECP_F3_4; 1.
 KM Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 963 AA; 108631 MW; 15A1A7CA933FB7B CRC64;
 Query Match 50.9%; Score 58; DB 2; Length 963;
 Best Local Similarity 47.8%; Pred. No. 15;
 Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
 QY 1 LPPVIVAVIGICAAVFLQQYV 23
 DB 646 SLVPTILAVVIGITLVAVTV 668
 RESULT 10
 Q6Z151.ORYSA PRELIMINARY; PRT; 162 AA.
 AC Q6Z151;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein P0443H10.13 (Hypothetical protein
 DE OSUNBa0026122.28).
 GN Name=P0443H10.13; Synonyms=OSUNBa0026122.28;
 OS Oryza sativa (Japanese cultivar-group).
 OC Bakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 CC EMBL; F0443H10.1; -!- Genomic DNA.
 CC NCBI_TaxID=39947;
 CC [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nuphobare(GA3) genomic DNA, chromosome 7, PAC
 RT clone:P0443H10.1";
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nuphobare(GA3) genomic DNA, chromosome 7, BAC
 RT clone:OSUNBa0026122.28";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005451; BAC84299.1; -!- Genomic DNA.
 DR EMBL; AP005101; BAC30841.1; -!- Genomic DNA.
 KM Gramine; O6Z151; -!
 KW Hypothetical protein.
 SQ SEQUENCE 162 AA; 17481 MW; 5590226D879F6872 CRC64;

Query Match 47.4%; Score 54; DB 2; Length 162;
 Best Local Similarity 55.6%; Pred. No. 12;
 Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 2 LPPVIVAVIGICAAVFL 19
 DB 15 LPPVIVAVISLPVAVFL 32
 RESULT 11
 MGRI_CABEL STANDARD; PRT; 999 AA.
 ID MGRI_CABEL
 AC 009630;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Probable metabotropic glutamate receptor mgl-1.
 GN Name=mgl-1; ORFNames=ZC506.4;
 OS Caenorhabditis elegans.
 OC Bakaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 CC NCBI_TaxID=6239;
 CC [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 CC -!- FUNCTION: Receptor for glutamate. The activity of this receptor is
 mediated by a G-protein that inhibits adenylate cyclase activity
 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 3 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; Z47073; CA87374.1; -!- Genomic DNA.
 DR PIR; T27628; T27628.
 DR HSSP; P23385; LEWT.
 DR Ensemble; ZC506.4; Caenorhabditis elegans.
 DR WormBase; WBGene0000332; mgl-1.
 DR WormPep; ZC506.4; CE01682.
 DR InterPro; IPR001828; ANF_rcgp_1lg_bd.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR InterPro; IPR000162; MrglR_receptor.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR Pfam; PF07562; NCD3G; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PRINTS; PR00593; MTABOTROPIC.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS50259; G_PROTEIN_RECP_F3_4; 1.
 KM Complete proteome; G-protein coupled receptor; Glycoprotein; Receptor;
 KW Transducer; Transmembrane.
 FT TRANSMEM 682 704 Potential.
 FT TRANSMEM 719 739 Potential.
 FT TRANSMEM 751 769 Potential.
 FT TRANSMEM 792 812 Potential.
 FT TRANSMEM 836 857 Potential.
 FT TRANSMEM 871 893 Potential.
 FT TRANSMEM 904 929 Potential.
 FT COMPIAS 888 891 Potential.
 FT CARBOHYD 518 518 Poly-Phe.
 SQ SEQUENCE 999 AA; 113276 MW; AB22AP2A28D9ACF CRC64;

Query Match 47.4%; Score 54; DB 1; Length 999;
 Best Local Similarity 47.8%; Pred. No. 57;
 Matches 11; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 NLPPIVAVTIGICAAVFLIQYV 23
 Db 682 SLVPIVAVTIGICAAVFLVIVVV 704

RESULT 12

Q5WCUG_BACSK PRELIMINARY; PRT; 278 AA.
 AC Q5WCUG;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Sugar ABC transporter permease.
 GN OrderedLocuNames=ABC3280;
 OS Bacillus clausii (strain KSM-K16).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=66692;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Takaki Y., Kageyama Y., Shimamura S., Suzuki H., Nishi S., Hatada Y.,
 RA Kawai S., Ito S., Horikoshi K.;
 RT "The complete genome sequence of the alkaliphilic Bacillus clausii
 KSM-K16."
 RT Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Part of a binding-protein-dependent transport system.
 CC Probably responsible for the translocation of the substrate across
 CC the membrane (By similarity).
 CC -1- SIMILARITY: Belongs to the binding membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport
 CC system permease family.
 DR EMBL; AP006627; BAB5814.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp_1; 1.
 DR PROSITE; PS50928; ABC_TM1; 1.
 DR Complete proteome; Transmembrane; Transport.
 SQ SEQUENCE 278 AA; 31819 MW; 32C334A7D874978 CRC64;

Query Match 46.5%; Score 53; DB 2; Length 278;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 11; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 NLPPIVAVTIGICAAVFLIQYV 22
 Db 141 NLPPIVAVTIGICAAVFLIRP 162

RESULT 13

O6AHM4_LEIXX PRELIMINARY; PRT; 285 AA.
 AC O6AHM4;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Rhomboid family membrane protein.
 GN Name=YGGP; OrderedLocuNames=LXK00170;
 OS Leifsonia xyli (subsp. xyli).
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Micrococciaceae; Microbacteriaceae; Leifsonia.
 OX NCBI_TaxID=59736;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CTC807;
 RX PubMed=15305603;
 RX Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,

RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,
 RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
 RA Taktia M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
 RA Almeida N.F., Jr., Carrier H., Coutinho L.L., El-Dorri H.A.,
 RA Perro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
 RA Goldmann G.H., Kimura E.T., Ferro E.S., Kuramae E.G.M.,
 RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.P.,
 RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
 RA Tsai S.M., Zanca A.S., Simpson A.J.G., Brumby S.M., Secubal J.C.;
 RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
 xyli subsp. xyli."
 RT Mol. Plant Microbe Interact. 17:827-836(2004).
 RL EMBL; AB016822; AAT8121.1; -; Genomic DNA.
 DR InterPro; IPR002610; Rhomboid_1like.
 DR Pfam; PF01694; Rhomboid; 1.
 KW Complete proteome.
 SQ SEQUENCE 285 AA; 30468 MW; CE117BE13912B0E CRC64;

Qy 3 PPIVAVTIGICAAVFLIQY 20
 Db 73 PPIVAVTIGICAAVFLIQY 90

RESULT 14

O6SN76_BACLD PRELIMINARY; PRT; 429 AA.
 AC O6SN76; Q62YM6;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
 DE Hypothetical protein (Na+/H+ antiporter).
 GN Name=nacA; OrderedLocuNames=BL05036, BL00533;
 OS Bacillus licheniformis (strain DSM 13 / ATCC 14580).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=279010;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15383718; DOI=10.1159/000079829;
 RA Velth B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
 RA Ehrenreich P., Baesner S., Henne A., Liesegang H., Merkl R.,
 RA Ehrenreich A., Gottschalk G.;
 RT "The complete genome sequence of Bacillus licheniformis DSM13, an
 RT organism with great industrial potential."
 RT J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;
 RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karlin S.D., Zaretsky E.J.,
 RA Tang M., Lopez de Leon A., Xiang H., Guet V., Clausen I.G.,
 RA Larsen F.B., Kasmsen M.D., Andersen J.T., Joergensen P.U.,
 RA Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Gatteron N.,
 RA Ehrlich S.D., Berka R.M.;
 RT "Complete genome sequence of the industrial bacterium Bacillus
 RT licheniformis and comparisons with closely related Bacillus species";
 RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
 DR EMBL; AB017333; AAU9448.1; -; Genomic DNA.
 DR EMBL; CP000002; AAU22132.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0015385; F:ion:hydrogen antiporter activity; IEA.
 DR GO; GO:0006885; P:regulation of pH; IEA.
 DR GO; GO:0006814; P:sodium ion transport; IEA.
 DR InterPro; IPR004770; Antiport_nhc.
 DR Pfam; PF03553; Na_H_antiporter; 2.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 429 AA; 44398 MW; C3F83D01AC7A9684 CRC64;

Query Match 46.5%; Score 53; DB 2; Length 429;
 Best Local Similarity 47.4%; Pred. No. 38;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 PIVLAVIGICAAVFLQQY 22
 ||:|||||:|:
 DB 407 PILIGVGICATIFRLPRF 425

RESULT 15

Q5YYB9 NOCPA PRELIMINARY; PRT; 107 AA.
 ID Q5YYB9 NOCPA PRELIMINARY; PRT; 107 AA.
 AC Q5YYB9;
 DT 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE Putative multi-drug efflux transporter.
 GN OrderedLocustNames=afa19760;
 OS Nocardia farcinica.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacteriaceae; Nocardiaceae; Nocardia.
 OX NCBI_TaxId=37329;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IFM 10152;
 RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
 RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
 RA Shiba T., Hattori M.;
 RT "The complete genomic sequence of Nocardia farcinica IFM 10152."
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
 DR EMBL; AP006618; BAD56822.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR InterPro; IPR000390; Smr.
 DR Pfam; PF00893; Multi_Drug_Res; 1.
 DR Complete proteome; Transmembrane.
 KW COMPLETE 107 AA; 10959 MW; C0F2E7B02C7A062F CRC64;
 SQ SEQUENCE

Query Match 45.64; Score 52; DB 2; Length 107;
 Best Local Similarity 45.04; Pred. No. 16;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 2 LPPIVLAVIGICAAVFLQQ 21
 ||:|||||:|:
 DB 29 LTPSIVVVGICAAVFLPSQ 48

Search completed: January 23, 2006, 09:31:14
 Job time : 62.7 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838a-8

Perfect score: 114

Sequence: 1 NLPPIVAVIGICAAVFLQQYV 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.Main:*
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/US10a_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/US10b_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	23	4	US-10-665-184-8
2	114	100.0	23	5	US-10-942-300-8
3	114	100.0	23	5	US-10-501-838a-8
4	114	100.0	30	4	US-10-665-184-30
5	114	100.0	30	5	US-10-942-300-30
6	114	100.0	30	5	US-10-501-838a-30
7	64	56.1	128	4	US-10-424-599-189871
8	58	50.9	381	3	US-09-934-455-180
9	58	50.9	381	4	US-10-225-068-150
10	58	50.9	381	4	US-10-374-780a-2534
11	58	50.9	381	5	US-10-225-068-150
12	58	50.9	383	4	US-10-437-963-188431
13	58	50.9	389	4	US-10-437-963-164930
14	57	50.0	345	4	US-10-425-114-71489
15	57	50.0	346	4	US-10-425-115-289943
16	56	49.1	104	4	US-10-767-701-51855
17	56	48.2	143	4	US-10-425-115-201325
18	55	48.1	234	4	US-10-424-599-266133
19	54	47.4	337	4	US-10-425-114-36852
20	54	47.4	367	4	US-10-424-599-150398
21	54	47.4	367	4	US-10-369-893-6871
22	53	46.5	1032	4	US-10-437-963-202452
23	52	45.6	375	3	US-09-934-455-474
24	51	44.7	289	4	US-10-437-963-195929
25	50.5	44.3	747	4	US-10-371-877-44
26	50	43.9	584	5	US-10-828-332-7
27	50	43.9	604	3	US-09-820-809-13

Applicant

28	50	43.9	983	4	US-10-255-149-2	Sequence 2, Appli
29	49.5	43.4	831	4	US-10-723-860-2937	Sequence 2937, Ap
30	49	43.0	23	5	US-10-665-184-2	Sequence 2, Appli
31	49	43.0	23	5	US-10-942-300-2	Sequence 2, Appli
32	49	43.0	23	5	US-10-501-838a-2	Sequence 2, Appli
33	49	43.0	25	4	US-10-665-184-13	Sequence 13, Appli
34	49	43.0	25	5	US-10-942-300-13	Sequence 13, Appli
35	49	43.0	25	5	US-10-501-838a-13	Sequence 13, Appli
36	49	43.0	179	4	US-10-223-047-11	Sequence 60, Appli
37	49	43.0	203	4	US-10-665-184-60	Sequence 60, Appli
38	49	43.0	203	5	US-10-942-300-60	Sequence 54, Appli
39	49	43.0	203	5	US-10-501-838a-54	Sequence 16, Appli
40	49	43.0	210	3	US-09-817-464-16	Sequence 16, Appli
41	49	43.0	210	4	US-10-331-289-16	Sequence 11, Appli
42	49	43.0	364	4	US-10-108-260A-4612	Sequence 4612, Ap
43	49	43.0	481	3	US-09-817-464-8	Sequence 8, Appli
44	49	43.0	481	4	US-10-331-289-8	Sequence 53626, A
45	49	43.0	619	4	US-10-282-122A-53626	

ALIGNMENTS

RESULT 1
US-10-665-184-8
Sequence 8, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
TITLE OF INVENTION: Biological Barrier
FILE REFERENCE: 24348-501CIP
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: US/10/665,184
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 8
LENGTH: 23
TYPE: PRT
ORGANISM: Rhizobium loti
US-10-665-184-8

Query Match 100.0%; Score 114; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.5e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLPPIVAVIGICAAVFLQQYV 23
DB 1 NLPPIVAVIGICAAVFLQQYV 23

RESULT 2
US-10-942-300-8
Sequence 8, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
TITLE OF INVENTION: Biological Barrier
FILE REFERENCE: 24348-503
CURRENT APPLICATION NUMBER: US/10/942,300
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Rhizobium lot1
US-10-942-300-8

Query Match 100.0%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.5e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLPPIVAVIGICAAVFLQQYV 23
Db 1 NLPPIVAVIGICAAVFLQQYV 23

RESULT 3
US-10-501-838A-8

; Sequence 8, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:

; APPLICANT: Ben-Sasson, Shmuel A.

; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

; FILE REFERENCE: 24348-501 NATL

; CURRENT APPLICATION NUMBER: US/10/501,838A

; PRIOR FILING DATE: 2004-07-19

; PRIOR APPLICATION NUMBER: PCT/IB03/00968

; PRIOR FILING DATE: 2003-02-07

; PRIOR APPLICATION NUMBER: US 60/355,396

; PRIOR FILING DATE: 2002-02-07

; NUMBER OF SEQ ID NOS: 72

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 8

; LENGTH: 23

; TYPE: PRT

; ORGANISM: Rhizobium lot1
US-10-501-838A-8

Query Match 100.0%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.5e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLPPIVAVIGICAAVFLQQYV 23
Db 1 NLPPIVAVIGICAAVFLQQYV 23

RESULT 4
US-10-665-184-30

; Sequence 30, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:

; APPLICANT: Cohen, Elnat

; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

; FILE REFERENCE: 24348-501CIP

; CURRENT APPLICATION NUMBER: US/10/665,184

; PRIOR FILING DATE: 2003-09-17

; PRIOR APPLICATION NUMBER: PCT/IB03/00968

; PRIOR FILING DATE: 2003-02-07

; PRIOR APPLICATION NUMBER: 60/355,396

; PRIOR FILING DATE: 2002-02-07

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30

; LENGTH: 30

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:
; OTHER INFORMATION: Penetrating peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)-(1)
; OTHER INFORMATION: ACETYLATION
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (30)-(30)
; OTHER INFORMATION: wherein Xaa is Lysine-NH2
US-10-665-184-30

Query Match 100.0%; Score 114; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLPPIVAVIGICAAVFLQQYV 23
Db 1 NLPPIVAVIGICAAVFLQQYV 23

RESULT 5
US-10-942-300-30

; Sequence 30, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:

; APPLICANT: Ben-Sasson, Shmuel

; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a

; FILE REFERENCE: 24348-503

; CURRENT APPLICATION NUMBER: US/10/942,300

; PRIOR FILING DATE: 2004-09-16

; PRIOR APPLICATION NUMBER: 10/665,184

; PRIOR FILING DATE: 2003-09-17

; PRIOR APPLICATION NUMBER: 10/664,989

; PRIOR FILING DATE: 2003-09-17

; PRIOR APPLICATION NUMBER: 60/503,615

; PRIOR FILING DATE: 2003-09-17

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30

; LENGTH: 30

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: Acylated Penetrating Peptide

; NAME/KEY: MISC_FEATURE

; LOCATION: (30)-(30)

; OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated

US-10-942-300-30

Query Match 100.0%; Score 114; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLPPIVAVIGICAAVFLQQYV 23
Db 1 NLPPIVAVIGICAAVFLQQYV 23

RESULT 6
US-10-501-838A-30

; Sequence 30, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:

; APPLICANT: Ben-Sasson, Shmuel A.

; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

; FILE REFERENCE: 24348-501 NATL

; CURRENT APPLICATION NUMBER: 10/665,184

; PRIOR FILING DATE: 2003-09-17

; PRIOR APPLICATION NUMBER: PCT/IB03/00968

; PRIOR FILING DATE: 2003-02-07

; PRIOR APPLICATION NUMBER: 60/355,396

; PRIOR FILING DATE: 2002-02-07

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 30

; LENGTH: 30

; TYPE: PRT

; ORGANISM: Artificial


```
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2003-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 30
/ LENGTH: 30
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic: penetrating peptide
/ NAME/KEY: MISC FEATURE
/ LOCATION: (27)..(27)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (30)..(30)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ NAME/KEY: MISC FEATURE
/ LOCATION: (30)..(30)
/ OTHER INFORMATION: wherein another molecule is coupled to the penetrating peptide
/ OTHER INFORMATION: via the free amino groups of the lysine residue
/ NAME/KEY: MISC FEATURE
/ LOCATION: (30)..(30)
/ OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
/ OTHER INFORMATION: peptide via the free amino groups of the lysine residue
US-10-501-838A-30
```

```
Query Match      100.0%; Score 114; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 NLPPIVAVIGICAVFLQQYV 23
DB      1 NLPPIVAVIGICAVFLQQYV 23
```

```
RESULT 7
US-10-424-599-189871
/ Sequence 189871, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 189871
/ LENGTH: 128
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_142470C.1.Pep
US-10-424-599-189871
```

```
Query Match      56.1%; Score 64; DB 4; Length 128;
Best Local Similarity 54.5%; Pred. No. 0.1;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 NLPPIVAVIGICAVFLQQY 22
```

```
DB      67 NFPPIVAVIGICAVFLVSY 88
```

```
RESULT 8
US-09-934-455-180
/ Sequence 180, Application US/09934455
/ Publication No. US20030121070A1
/ GENERAL INFORMATION:
/ APPLICANT: Adam, Luc
/ APPLICANT: Creelman, Robert
/ APPLICANT: Dubell, Arnold
/ APPLICANT: Heard, Jacqueline
/ APPLICANT: Jiang, Cai-Zhong
/ APPLICANT: Keddie, James
/ APPLICANT: Pilgrim, Marsha
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Reuber, Lynne
/ APPLICANT: Riechmann, Jose Luis
/ APPLICANT: Yu, Guo-Liang
/ APPLICANT: Pineda, Omaira
/ TITLE OF INVENTION: Genes for Modifying Plant Traits IV
/ FILE REFERENCE: MBI-0025
/ CURRENT APPLICATION NUMBER: US/09/934,455
/ CURRENT FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 60/227439
/ PRIOR FILING DATE: 2000-08-22
/ PRIOR APPLICATION NUMBER: MBI-0022
/ PRIOR FILING DATE: 2001-11-16
/ PRIOR APPLICATION NUMBER: MBI-0023
/ PRIOR FILING DATE: 2001-04-17
/ NUMBER OF SEQ ID NOS: 516
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 180
/ LENGTH: 381
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
US-09-934-455-180
```

```
Query Match      50.9%; Score 58; DB 3; Length 381;
Best Local Similarity 54.5%; Pred. No. 2.7;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 NFPPIVAVIGICAVFLQY 22
DB      40 NFPPIVAVIGICAVFLVSY 61
```

```
RESULT 9
US-10-225-068-150
/ Sequence 150, Application US/10225068
/ Publication No. US20030217383A1
/ GENERAL INFORMATION:
/ APPLICANT: Mendel Biotechnology, Inc.
/ APPLICANT: Reuber, T. Lynne
/ APPLICANT: Riechmann, Jose Luis
/ APPLICANT: Heard, Jacqueline E.
/ APPLICANT: Jiang, Cai-Zhong
/ APPLICANT: Adam, Luc J.
/ APPLICANT: Dubell, Arnold T.
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Pineda, Omaira
/ APPLICANT: Yu, Guo-Liang
/ APPLICANT: Broun, Pierre E.
/ TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
/ TITLE OF INVENTION: POLYPEPTIDES IN PLANTS
/ FILE REFERENCE: 514442002040
/ CURRENT APPLICATION NUMBER: US/10/225,068
/ CURRENT FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 60/310,847
/ PRIOR FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: 60/336,049
/ PRIOR FILING DATE: 2001-11-19
```

```
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (134)...(159)
; OTHER INFORMATION: Conserved domain
US-10-225-068-150

Query Match      50.9%; Score 58; DB 4; Length 381;
Best Local Similarity 54.5%; Pred. No. 2.7;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy      1 NLPPIVAVIGICAVFLQQY 22
Db      40 NFPIATAVIGILATAPFLVST 61

RESULT 10
; Sequence 2534, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2534
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
```

```
; FEATURE:
; OTHER INFORMATION: G1095
US-10-374-780A-2534

Query Match      50.9%; Score 58; DB 4; Length 381;
Best Local Similarity 54.5%; Pred. No. 2.7;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy      1 NLPPIVAVIGICAVFLQQY 22
Db      40 NFPIATAVIGILATAPFLVSY 61

RESULT 11
US-10-225-068-150
; Sequence 150, Application US/10225068
; Publication No. US20050120408A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Adam, Luc J.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Brown, Pierre E
; TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
; FILE REFERENCE: 51442002040
; CURRENT APPLICATION NUMBER: US/10/225,068
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (134)...(159)
; OTHER INFORMATION: Conserved domain
US-10-225-068-150

Query Match      50.9%; Score 58; DB 5; Length 381;
Best Local Similarity 54.5%; Pred. No. 2.7;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy      1 NLPPIVAVIGICAVFLQQY 22
Db      40 NFPIATAVIGILATAPFLVSY 61

RESULT 12
US-10-437-963-188431
; Sequence 188431, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
```

```

/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 188431
/ LENGTH: 383
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_85035C.1.pep
US-10-437-963-188431

Query Match          50.9%; Score 58; DB 4; Length 383;
Best Local Similarity 52.6%; Pred. No. 2.7;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 PIVAVIGICAAVFLQY 22
DB 57 PLVIAITIGLASAPFLVSY 75

RESULT 13
US-10-437-963-164930
/ Sequence 164930, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 164930
/ LENGTH: 389
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_63783C.1.pep
US-10-437-963-164930

Query Match          50.9%; Score 58; DB 4; Length 389;
Best Local Similarity 52.6%; Pred. No. 2.7;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 PIVAVIGICAAVFLQY 22
DB 70 PLVIAITIGLASAPFLVSY 88

RESULT 14
US-10-425-114-71489
/ Sequence 71489, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei

```

```

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 71489
/ LENGTH: 345
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMROB73017C10_PFI.pep
US-10-425-114-71489

Query Match          50.0%; Score 57; DB 4; Length 345;
Best Local Similarity 52.6%; Pred. No. 3.4;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 PIVAVIGICAAVFLQY 22
DB 52 PLVIAITIGLASAPFLVSY 70

RESULT 15
US-10-425-115-289943
/ Sequence 289943, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 289943
/ LENGTH: 346
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_27515C.1.pep
US-10-425-115-289943

Query Match          50.0%; Score 57; DB 4; Length 346;
Best Local Similarity 52.6%; Pred. No. 3.4;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 PIVAVIGICAAVFLQY 22
DB 52 PLVIAITIGLASAPFLVSY 70

Search completed: January 23, 2006, 11:52:34
Job time : 60 secs

```

This page blank (upside)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:56 ; Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838a-9

Perfect score: 119

Sequence: 1 NYFIVNLALADLCMAAFNAFNF 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pap:*
2: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pap:*
3: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pap:*
4: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pap:*
5: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pap:*
6: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pap:*
7: /cgn2_6/prodata/2/pubppa/US11_NEW_PUB.pap:*
8: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	66.4	407	7	US-11-127-877-56
2	66	55.5	23	7	US-11-127-877-571
3	65	54.6	259	7	US-11-151-482-5
4	64	45.4	429	7	US-11-127-877-51
5	54	45.4	466	7	US-11-127-877-50
6	51	42.9	409	6	US-10-627-633-4
7	51	42.9	415	6	US-10-627-633-2
8	51	42.9	445	7	US-11-115-564-1
9	51	42.9	445	7	US-11-115-564-2
10	51	42.9	445	7	US-11-115-564-3
11	51	42.9	461	7	US-11-080-439-16
12	50	42.0	95	6	US-10-521-162-37
13	50	42.0	359	6	US-10-995-561-712
14	50	42.0	359	6	US-10-995-561-716
15	50	42.0	359	6	US-10-876-787-2
16	50	42.0	359	7	US-11-127-877-65
17	50	42.0	388	6	US-10-995-561-713
18	50	42.0	394	6	US-10-995-561-714
19	50	42.0	394	6	US-10-995-561-715
20	50	42.0	417	6	US-10-992-577-44
21	50	42.0	480	6	US-10-521-162-40
22	49	41.2	23	7	US-11-127-877-556
23	49	41.2	353	6	US-10-875-716-10
24	49	41.2	466	7	US-11-127-877-41
25	48	40.3	30	6	US-10-901-576-10

25	48	40.3	31	6	US-10-901-576-9	Sequence 9, Appl1
27	48	40.3	420	6	US-10-992-577-6	Sequence 6, Appl1
28	48	40.3	522	6	US-10-510-018-2	Sequence 2, Appl1
29	47	39.5	518	7	US-11-127-877-38	Sequence 38, Appl1
30	47	39.5	321	6	US-10-055-877-338	Sequence 338, Appl1
31	47	39.5	347	6	US-10-131-826A-18	Sequence 18, Appl1
32	47	39.5	387	6	US-10-467-657-5522	Sequence 5522, Appl1
33	47	39.5	430	6	US-10-992-577-8	Sequence 8, Appl1
34	47	39.5	432	6	US-10-992-577-2	Sequence 2, Appl1
35	47	39.5	532	7	US-11-127-877-42	Sequence 42, Appl1
36	46	38.7	271	6	US-10-980-388-67	Sequence 67, Appl1
37	46	38.7	471	6	US-10-995-561-901	Sequence 901, Appl1
38	45	37.8	120	6	US-10-055-877-179	Sequence 179, Appl1
39	45	37.8	259	6	US-10-055-877-225	Sequence 225, Appl1
40	45	37.8	259	6	US-10-055-877-237	Sequence 237, Appl1
41	45	37.8	352	6	US-10-627-633-6	Sequence 6, Appl1
42	45	37.8	384	7	US-11-080-991-26	Sequence 26, Appl1
43	45	37.8	590	7	US-11-124-368A-183	Sequence 183, Appl1
44	45	37.8	590	7	US-11-127-877-54	Sequence 54, Appl1
45	44	37.0	313	7	US-11-095-093-2	Sequence 2, Appl1

ALIGNMENTS

```

RESULT 1
US-11-127-877-56
; Sequence 56, Application US/11127877
; Publication No. US20050287565A1
; GENERAL INFORMATION:
; APPLICANT: Merchiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koentraad F. P.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Method, Compositions and Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27, 800-B USA
; CURRENT APPLICATION NUMBER: US/11/127,877
; PRIOR FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; SOFTWARE: PatentIn version 3.3
; NUMBER OF SEQ ID NOS: 590
; SEQ ID NO 56
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-877-56

Query Match      66.4%; Score 79; DB 7; Length 407;
Best Local Similarity 65.2%; Pred. No. 1.7e-05;
Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY      1 NYFIVNLALADLCMAAFNAFNF 23
Db      68 NYFIVNLALADLCMAAFNAFNF 90

RESULT 2
US-11-127-877-571
; Sequence 571, Application US/11127877
; Publication No. US20050287565A1
; GENERAL INFORMATION:
; APPLICANT: Merchiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koentraad F. P.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Method, Compositions and Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27, 800-B USA
; CURRENT APPLICATION NUMBER: US/11/127,877

```

```
/ CURRENT FILING DATE: 2005-05-12
/ PRIOR APPLICATION NUMBER: 60/570,352
/ PRIOR FILING DATE: 2004-05-12
/ PRIOR APPLICATION NUMBER: 60/603,948
/ PRIOR FILING DATE: 2004-08-24
/ NUMBER OF SEQ ID NOS: 590
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 571
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-127-877-571

Query Match          55.5%; Score 66; DB 7; Length 23;
Best Local Similarity 61.9%; Pred. No. 0.0001;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy          3 FIVNLALADLCMAAFNAAF 23
            :|:|:|:|:|:|:|:|:|:|
Db          1 FLVNLAPASWAAFTVTVNF 21

RESULT 3
US-11-151-482-5
/ Sequence 5, Application US/1151482
/ Publication No. US2006002919A1
/ GENERAL INFORMATION:
/ APPLICANT: Glucksmann, Maria A.
/ APPLICANT: Gu, Wei
/ TITLE OF INVENTION: 15625 Receptor, A Novel G-Protein Coupled Receptor
/ FILE REFERENCE: 5800-13, 035800-171548
/ CURRENT APPLICATION NUMBER: US/11/151,482
/ CURRENT FILING DATE: 2005-06-13
/ PRIOR APPLICATION NUMBER: US/09/187,134
/ PRIOR FILING DATE: 1998-11-06
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 269
/ TYPE: PRT
/ ORGANISM: Unknown
/ FEATURE:
/ OTHER INFORMATION: Description of Unknown Organism: Rhodopsin family
US-11-151-482-5

Query Match          54.6%; Score 65; DB 7; Length 269;
Best Local Similarity 61.9%; Pred. No. 0.0019;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy          1 NYFIVNLALADLCMAAFNAAF 21
            :|:|:|:|:|:|:|:|:|:|
Db          20 NYFIVNLAVADLFLSLFTMPF 40

RESULT 4
US-11-127-877-51
/ Sequence 51, Application US/11127877
/ Publication No. US20050287565A1
/ GENERAL INFORMATION:
/ APPLICANT: Merchiers, Pascal G.
/ APPLICANT: Hoffmann, Marcel
/ APPLICANT: Spittaels, Koenraad F. F.
/ APPLICANT: Laenen, Wendy
/ TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
/ TITLE OF INVENTION: Amyloid-Beta Protein Production
/ FILE REFERENCE: P27,800-B USA
/ CURRENT APPLICATION NUMBER: US/11/127,877
/ CURRENT FILING DATE: 2005-05-12
/ PRIOR APPLICATION NUMBER: 60/570,352
/ PRIOR FILING DATE: 2004-05-12
/ PRIOR APPLICATION NUMBER: 60/603,948
/ PRIOR FILING DATE: 2004-08-24
```

```
/ NUMBER OF SEQ ID NOS: 590
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 51
/ LENGTH: 429
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-127-877-51

Query Match          45.4%; Score 54; DB 7; Length 429;
Best Local Similarity 48.0%; Pred. No. 0.17;
Matches 12; Conservative 6; Mismatches 3; Indels 4; Gaps 1;

Qy          1 NYFIVNLALADLCMAA----FNAAF 21
            :|:|:|:|:|:|:|:|:|:|
Db          62 HYIVNLAVADLLTSTVLPFSALF 86

RESULT 5
US-11-127-877-50
/ Sequence 50, Application US/11127877
/ Publication No. US20050287565A1
/ GENERAL INFORMATION:
/ APPLICANT: Merchiers, Pascal G.
/ APPLICANT: Hoffmann, Marcel
/ APPLICANT: Spittaels, Koenraad F. F.
/ APPLICANT: Laenen, Wendy
/ TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
/ TITLE OF INVENTION: Amyloid-Beta Protein Production
/ FILE REFERENCE: P27,800-B USA
/ CURRENT APPLICATION NUMBER: US/11/127,877
/ CURRENT FILING DATE: 2005-05-12
/ PRIOR APPLICATION NUMBER: 60/570,352
/ PRIOR FILING DATE: 2004-05-12
/ PRIOR APPLICATION NUMBER: 60/603,948
/ PRIOR FILING DATE: 2004-08-24
/ NUMBER OF SEQ ID NOS: 590
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 50
/ LENGTH: 466
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-127-877-50

Query Match          45.4%; Score 54; DB 7; Length 466;
Best Local Similarity 48.0%; Pred. No. 0.18;
Matches 12; Conservative 6; Mismatches 3; Indels 4; Gaps 1;

Qy          1 NYFIVNLALADLCMAA----FNAAF 21
            :|:|:|:|:|:|:|:|:|:|
Db          62 HYIVNLAVADLLTSTVLPFSALF 86

RESULT 6
US-10-627-633-4
/ Sequence 4, Application US/10627633
/ Publication No. US20050250720A1
/ GENERAL INFORMATION:
/ APPLICANT: Charles, Andrew David
/ APPLICANT: Brennard, John Charles
/ APPLICANT: Hart, Kevin Anthony
/ TITLE OF INVENTION: Novel Compound
/ FILE REFERENCE: 1991-221
/ CURRENT APPLICATION NUMBER: US/10/627,633
/ CURRENT FILING DATE: 2003-07-28
/ PRIOR APPLICATION NUMBER: 09/722,342
/ PRIOR FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: 60/172,146
/ PRIOR FILING DATE: 1999-12-17
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 4
/ LENGTH: 409
/ TYPE: PRT
```

ORGANISM: Mus musculus
US-10-627-633-4

Query Match 42.9%; Score 51; DB 6; Length 409;
Best Local Similarity 42.9%; Pred. No. 0.47;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 NYFVNLALADLCMAAF 21
DB 95 NYFVNSMACADLISVASTPF 115

RESULT 7
US-10-627-633-2
Sequence 2, Application US/10627633
Publication No. US20050250720A1
GENERAL INFORMATION:
APPLICANT: Charles, Andrew David
APPLICANT: Brennan, John Charles
APPLICANT: Hart, Kevin Anthony
TITLE OF INVENTION: Novel Compound
FILE REFERENCE: 1991-221
CURRENT APPLICATION NUMBER: US/10/627,633
PRIOR FILING DATE: 2003-07-28
PRIOR APPLICATION NUMBER: 09/722,342
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 60/172,146
PRIOR FILING DATE: 1999-12-17
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin version 3.2
SEQ ID NO 2
LENGTH: 415
TYPE: PRT
ORGANISM: Homo sapiens
US-10-627-633-2

Query Match 42.9%; Score 51; DB 6; Length 415;
Best Local Similarity 42.9%; Pred. No. 0.48;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 NYFVNLALADLCMAAF 21
DB 101 NYFVNSMACADLISVASTPF 121

RESULT 8
US-11-115-564-1
Sequence 1, Application US/11115564
Publication No. US20050267116A1
GENERAL INFORMATION:
APPLICANT: Peschke, Bernd
APPLICANT: Hohlweg, Rolf
TITLE OF INVENTION: SUBSTITUTED HEXAHYDROPYRROLO[1,2-A]PYRAZINES,
TITLE OF INVENTION: OCTAHYDROPYRIDO[1,2-A]PYRAZINES AND
FILE REFERENCE: 6483,200-US
CURRENT APPLICATION NUMBER: US/11/115,564
PRIOR FILING DATE: 2005-04-27
PRIOR APPLICATION NUMBER: US 60/387,047
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: Danish Application no. PA 2002 00863
PRIOR FILING DATE: 2002-06-06
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin version 3.2
SEQ ID NO 1
LENGTH: 445
TYPE: PRT
ORGANISM: Homo Sapiens
US-11-115-564-1

Query Match 42.9%; Score 51; DB 7; Length 445;
Best Local Similarity 47.1%; Pred. No. 0.52;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYFVNLALADLCMAAF 17
DB 70 NFPLNLALSDPLVGAF 86

RESULT 9
US-11-115-564-2
Sequence 2, Application US/11115564
Publication No. US20050267116A1
GENERAL INFORMATION:
APPLICANT: Peschke, Bernd
APPLICANT: Hohlweg, Rolf
TITLE OF INVENTION: SUBSTITUTED HEXAHYDROPYRROLO[1,2-A]PYRAZINES,
TITLE OF INVENTION: OCTAHYDROPYRIDO[1,2-A]PYRAZINES AND
FILE REFERENCE: 6483,200-US
CURRENT APPLICATION NUMBER: US/11/115,564
PRIOR FILING DATE: 2005-04-27
PRIOR APPLICATION NUMBER: US 60/387,047
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: Danish Application no. PA 2002 00863
PRIOR FILING DATE: 2002-06-06
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin version 3.2
SEQ ID NO 2
LENGTH: 445
TYPE: PRT
ORGANISM: Monkey
US-11-115-564-2

Query Match 42.9%; Score 51; DB 7; Length 445;
Best Local Similarity 47.1%; Pred. No. 0.52;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYFVNLALADLCMAAF 17
DB 70 NFPLNLALSDPLVGAF 86

RESULT 10
US-11-115-564-3
Sequence 3, Application US/11115564
Publication No. US20050267116A1
GENERAL INFORMATION:
APPLICANT: Peschke, Bernd
APPLICANT: Hohlweg, Rolf
TITLE OF INVENTION: SUBSTITUTED HEXAHYDROPYRROLO[1,2-A]PYRAZINES,
TITLE OF INVENTION: OCTAHYDROPYRIDO[1,2-A]PYRAZINES AND
FILE REFERENCE: 6483,200-US
CURRENT APPLICATION NUMBER: US/11/115,564
PRIOR FILING DATE: 2005-04-27
PRIOR APPLICATION NUMBER: US 60/387,047
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: Danish Application no. PA 2002 00863
PRIOR FILING DATE: 2002-06-06
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin version 3.2
SEQ ID NO 3
LENGTH: 445
TYPE: PRT
ORGANISM: Rat
US-11-115-564-3

Query Match 42.9%; Score 51; DB 7; Length 445;
Best Local Similarity 47.1%; Pred. No. 0.52;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYFVNLALADLCMAAF 17
DB 70 NFPLNLALSDPLVGAF 86

```
RESULT 11
US-11-090-439-16
; Sequence 16, Application US/11090439
; Publication No. US2005026442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-16

Query Match          42.9%; Score 51; DB 7; Length 481;
Best Local Similarity 52.9%; Pred. No. 0.56;
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYFVNLADLCMAAF 17
|:::|||||:
Db 90 NYFVSLAVADLVGLF 106

RESULT 12
US-10-521-162-37
; Sequence 37, Application US/10521162
; Publication No. US20050287529A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: FLRA AND TICK OCTOPAMINE RECEPTOR NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-11-PCT
; CURRENT APPLICATION NUMBER: US/10/521,162
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: 60/319,402
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/426,601
; PRIOR FILING DATE: 2003-11-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Rhinipicephalus sanguineus
US-10-521-162-37

Query Match          42.0%; Score 50; DB 6; Length 95;
Best Local Similarity 75.0%; Pred. No. 0.15;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFVNLADLC 12
|:::|||||:
Db 60 NYFVSLAVADLV 71

RESULT 13
US-10-995-561-712
; Sequence 712, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
```

```
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 712
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-712

Query Match          42.0%; Score 50; DB 6; Length 359;
Best Local Similarity 81.8%; Pred. No. 0.6;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 FIVNLADLC 13
|:::|||||:
Db 66 FIVNLADLC 76

RESULT 14
US-10-995-561-716
; Sequence 716, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 716
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-716

Query Match          42.0%; Score 50; DB 6; Length 359;
Best Local Similarity 81.8%; Pred. No. 0.6;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 FIVNLADLC 13
|:::|||||:
Db 66 FIVNLADLC 76

RESULT 15
US-10-876-787-2
; Sequence 2, Application US/10876787
; Publication No. US20050287535A1
; GENERAL INFORMATION:
; APPLICANT: McGrath, Kevin P.
; APPLICANT: Kimberly-Clark Worldwide, Inc.
; TITLE OF INVENTION: BIOMARKERS FOR WOUND HEALING
; FILE REFERENCE: 1443.156US1
; CURRENT APPLICATION NUMBER: US/10/876,787
; CURRENT FILING DATE: 2004-06-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-876-787-2

Query Match          42.0%; Score 50; DB 6; Length 359;
Best Local Similarity 81.8%; Pred. No. 0.6;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 FIVNLADLC 13
```


Db 66 PLINIALADLC 76

Search completed: January 23, 2006, 11:53:36
Job time : 5.1 secs

This page blank (uspo)
(Circin) kunoj e60d 341

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-9

Perfect score: 119
Sequence: 1 NTFIVNLALADICMAAFNAFNAF 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_21.*
1: geneseqp19808.*
2: geneseqp19908.*
3: geneseqp20008.*
4: geneseqp20018.*
5: geneseqp20028.*
6: geneseqp20038.*
7: geneseqp20038.*
8: geneseqp20048.*
9: geneseqp20058.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	23	6 ADB16894	Adb16894 Human neu
2	119	100.0	23	9 AEB08226	Aeb08226 Human NK-
3	119	100.0	30	6 ADB16914	Adb16914 Escherich
4	119	100.0	30	9 AEB08248	Aeb08248 Penetrati
5	119	100.0	301	8 ADM72122	Adm72122 Human NTR
6	119	100.0	333	8 ADL70533	Adl70533 Human G-P
7	119	100.0	333	8 ABM83530	Abm83530 Human dia
8	119	100.0	385	1 AAP90551	Aap90551 Bovine su
9	119	100.0	387	7 ABU61827	Abu61827 Human sub
10	119	100.0	398	2 AAB80135	Aab80135 Human rec
11	119	100.0	398	4 AAB56367	Aab56367 Non-endog
12	119	100.0	398	5 AEG32383	Aeg32383 Human tac
13	119	100.0	398	5 AEG32382	Aeg32382 Human tac
14	119	100.0	398	6 ABP81985	Abp81985 Human tac
15	119	100.0	398	7 ADC86195	Adc86195 Human GPC
16	119	100.0	399	8 ADC29658	Adc29658 Human GPC
17	119	100.0	411	7 ADB08665	Adb08665 Novel pro
18	119	100.0	411	7 AAU08337	Aau08337 Bovine su
19	117	98.3	313	4 ABG73541	Abg73541 Bovine su
20	117	98.3	313	6 ABU62487	Abu62487 Bovine su
21	114	95.8	369	4 AAB66622	Aab66622 Rat subet
22	114	95.8	384	8 ADC29659	Adc29659 Mouse GPC
23	89	74.8	372	4 AAB66623	Aab66623 Rat subet
24	89	74.8	407	2 AAR12970	Aar12970 Substance

25	89	74.8	407	2 AAR32800	Aar32800 Rat Subst
26	89	74.8	407	8 ADC29657	Adc29657 Mouse GPC
27	79	66.4	23	6 ADB16907	Adb16907 Human neu
28	79	66.4	23	9 AEB08241	Aeb08241 Human NK-
29	79	66.4	29	6 ADB16917	Adb16917 Escherich
30	79	66.4	29	9 AEB08251	Aeb08251 Penetrati
31	79	66.4	311	2 AAR28800	Aar28800 Human neu
32	79	66.4	311	5 AAU91342	Aau91342 Neurokin
33	79	66.4	407	2 AAR27524	Aar27524 Human rec
34	79	66.4	407	2 AAR25306	Aar25306 hepr5urf2
35	79	66.4	407	2 AAR32799	Aar32799 Substance
36	79	66.4	407	4 ABB56366	Abb56366 Non-endog
37	79	66.4	407	5 ABB76988	Abb76988 Human tac
38	79	66.4	407	5 ABB76987	Abb76987 Human tac
39	79	66.4	407	5 ABB76990	Abb76990 Human tac
40	79	66.4	407	5 ABB76989	Abb76989 Human tac
41	79	66.4	407	5 ABB77000	Abb77000 Human tac
42	79	66.4	407	5 AAU91345	Aau91345 Neurokin
43	79	66.4	407	5 AAU91344	Aau91344 Neurokin
44	79	66.4	407	5 AAU78238	Aau78238 Human Tac
45	79	66.4	407	5 ABB99006	Abb99006 Mutant hu

ALIGNMENTS

RESULT 1
ID ADB16894 standard; peptide; 23 AA.
XX ADB16894;

20-NOV-2003 (first entry)

Human neurokinin 2 (NK-2) receptor penetrating peptide 9.

penetrating peptide; epithelial; endocheial; tight junction; diabetes;
infertility; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
osteopathic; cytostatic; nootropic.

Homo sapiens.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-1B000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating or preventing e.g. endocrine disorders.

Claim 2; Page 14; 60pp; English.

This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from the human
 CC neurokinin 2 (NK-2) receptor and is penetrating peptide 9 of the
 CC invention.

XX SQ Sequence 23 AA;

Query Match 100.0%; Score 119; DB 6; Length 23;

Best Local Similarity 100.0%; Pred. No. 2,3e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIVNLALADLCMAAFNAFNF 23
 |||||
 DB 1 NYFIVNLALADLCMAAFNAFNF 23

RESULT 2
 AEB08226

ID AEB08226 standard; peptide; 23 AA.

AC AEB08226;

XX 25-AUG-2005 (first entry)

DE Human NK-2 receptor penetrating peptide 9, SEQ ID NO: 9.

KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; arteriosclerosis;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antiandemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection; receptor;
 KW Natural killer 2; NK2.

XX Homo sapiens.

OS US2005136103-A1.

PN 23-JUN-2005.

PD 16-SEP-2004; 2004US-00942300.

PF 17-SEP-2003; 2003US-00664989.

PR 17-SEP-2003; 2003US-00665184.

PR 17-SEP-2003; 2003US-0503615P.

XX (BENS/) BEN-SASSON S A.

PA (COHE/) COHEN E.

XX Ben-Sasson SA, Cohen E;
 DR WPI, 2005-444089/45.

PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.

PS Claim 55, SEQ ID NO 9, 59pp; English.

CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the human NK-2 penetrating peptide.
 CC This sequence is used in the effective translocation of aminoglycoside
 CC antibiotics and antifungal agents across an epithelial barrier.

XX SQ Sequence 23 AA;

Query Match 100.0%; Score 119; DB 9; Length 23;

Best Local Similarity 100.0%; Pred. No. 2,3e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIVNLALADLCMAAFNAFNF 23
 |||||
 DB 1 NYFIVNLALADLCMAAFNAFNF 23

RESULT 3
 ADB16914

ID ADB16914 standard; peptide; 30 AA.

AC ADB16914;

XX 20-NOV-2003 (first entry)

DE Escherichia coli peptide 3 coupled to imaging compound linker, IBW-004.

KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiac; antiarteriosclerotic; osteopathic; cytostatic; neurotropic;
 KW imaging linker; penetrating peptide; IBW-004.

XX Synthetic.

OS Escherichia coli.

OS Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 30 /note= "C-terminal amide"

XX WO2003066859-A2.

XX 14-AUG-2003.

PP 07-FEB-2003; 2003WO-IB000968.

XX 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 XX WPI, 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX
 XX Example 3, Page 40, 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, haematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX Accordingly, the peptides of this invention can be used in compositions
 XX that have neuroprotective, cardiant, antihypertensive, osteopathic,
 XX cytostatic or neurotropic activities. This peptide sequence is IBM-004, to
 XX which consists of the Escherichia coli penetrating peptide 3 coupled to
 XX the imaging linker peptide used in an exemplification of the invention.
 XX
 XX Sequence 30 AA:
 XX
 XX Query Match 100.0%; Score 119; DB 6; Length 30;
 XX Best Local Similarity 100.0%; Pred. No. 3e-12;
 XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 NYPIVNLALADLCMAAFNAFNF 23
 XX Db 1 NYPIVNLALADLCMAAFNAFNF 23
 XX
 XX RESULT 4
 XX AEB08248
 XX ID AEB08248 standard; peptide; 30 AA.
 XX AC AEB08248;
 XX XX
 XX DT 25-AUG-2005 (first entry)
 XX DE Penetrating peptide SEQ: 31 used in composition for mucosal vaccination.
 XX XX
 XX XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 XX XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 XX XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 XX XX demencia; genetic disorder; cardiovascular disease;
 XX XX cardiovascular-gen.; atherosclerosis; antiarteriosclerosis;
 XX XX coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 XX XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX XX genitourinary disease; hematological disease; anemiaemic; anemia;
 XX XX autoimmune disease; immunosuppressive; immune deficiency; infection;
 XX XX immunostimulant; infectious disease; antimicrobial; infection;
 XX XX erectile dysfunction; andrology; major depressive disorder;
 XX XX antidepressant; psychiatric disorder; pain; analgesic;
 XX XX bacterial infection; antibacterial; viral infection; virucide;
 XX XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX XX renal failure; antifertility; antirheumatic; cytostatic;

KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX OS Unidentified.
 XX
 XX Key Location/Qualifiers
 XX FT Misc-difference 1
 XX FT Misc-difference 1 /note="N-terminal acylated"
 XX FT Misc-difference 30 /note="Optionally C-terminal amide, optionally the free
 XX FT amino group of lysine is acylated with a fatty acid"
 XX
 XX US2005136103-A1.
 XX
 XX PD 23-JUN-2005.
 XX
 XX PF 16-SEP-2004; 2004US-00942300.
 XX
 XX PR 17-SEP-2003; 2003US-00664989.
 XX PR 17-SEP-2003; 2003US-00665184.
 XX PR 17-SEP-2003; 2003US-0503615P.
 XX
 XX PA (BENS/) BEN-SASSON S A.
 XX PI (COHE/) COHEN E.
 XX
 XX PI Ben-Sasson SA, Cohen E;
 XX WPI, 2005-444089/45.
 XX
 XX Claim 63; SEQ ID NO 31; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for trans epithelial delivery of effector.
 XX The invention comprises the effector sequentially coupled with a counter ion
 XX and at least one hydrophobic agent, where the effector is selectively
 XX encapsulated into a complex. The invention is useful for translocating
 XX effectors across a biological barrier such as epithelial cells and
 XX endothelial cells during treatment and prevention of disease or
 XX pathological conditions (including endocrine disorders, diabetes,
 XX infertility, hormone deficiencies, osteoporosis, ophthalmological
 XX disorders, neurodegenerative disorders, Huntington's disease,
 XX Parkinson's disease, multiple sclerosis, hyper-coagulable states, hypo-
 XX cardiovascular disorders, atherosclerosis, cerebrovascular events, metabolic
 XX coagulable states, coronary disease, cerebrovascular events, renal failure,
 XX hematology disorders, anemia of different entities, immunologic and
 XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
 XX infectious diseases, viral infections, bacterial infections, fungal
 XX infections, parasitic infections, neoplastic diseases, multi-factorial
 XX disorders, impotence, chronic pain, depression, different fibrosis states
 XX and short stature) and for mucosal vaccination against anthrax and
 XX hepatitis B. The present sequence is a penetrating peptide (IBM-004) used
 XX in the composition for mucosal vaccination using a counter anion and a
 XX penetrating peptide.
 XX
 XX Sequence 30 AA:
 XX
 XX Query Match 100.0%; Score 119; DB 9; Length 30;
 XX Best Local Similarity 100.0%; Pred. No. 3e-12;
 XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 NYPIVNLALADLCMAAFNAFNF 23
 XX Db 1 NYPIVNLALADLCMAAFNAFNF 23
 XX
 XX RESULT 5
 XX ADM72122
 XX ID ADM72122 standard; protein; 301 AA.
 XX XX

AC ADM21212;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human NTRAN polypeptide (clone ID 7515474CD1).
XX
KW NTRAN; neurotransmission-associated protein; cyostatic;
KW antidiabetic; anti-HIV; antiallergic; cerebroprotective;
KW antiparkinsonian; anticonvulsant; dermatological; neurotrophic;
KW neuroprotective; antiinflammatory; antithyroid; antiarthritic; anorectic;
KW cardiac; hypotensive; hepatotropic; gene therapy; human.
XX
OS Homo sapiens.
XX
PM MO2004022705-A2.
XX
PD 18-MAR-2004.
XX
PF 02-SEP-2003; 2003WO-US027411.
XX
PR 04-SEP-2002; 2002US-0408383P.
PR 05-SEP-2002; 2002US-0408781P.
PR 26-SEP-2002; 2002US-041421P.
PR 13-NOV-2002; 2002US-0426483P.
PR 06-DEC-2002; 2002US-0431566P.
PR 17-DEC-2002; 2002US-0434317P.
PR 03-JAN-2003; 2003US-0437763P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Lee SY, Elliott VS, Hafalia AJA, Burford N, Sprague MW;
PI Griffin JB, Yang YG, Chawia NK, Baughn MR, Becha SD, Khare R;
PI Thornton MA, Mason PM, Gietzen KJ, Ison CH, Margulis JP, Swarnakar A;
PI Ramkumar J, Jin F, Richardson TW, Tran UK;
XX
XX WPI; 2004-269571/25.
DR N-Psdb; ADM72152.
XX
PT New human neurotransmission-associated proteins and polynucleotides for
PT diagnosing, preventing or treating diseases or conditions associated with
PT aberrant protein expression, e.g. cancer, hepatitis, AIDS, obesity or
PT stroke.
XX
PT
XX
PS Claim 1; SEQ ID NO 3; 216bp; English.
XX
XX
XX The invention relates to human neurotransmission-associated proteins
XX (NTRAN) and encoding polynucleotides. The NTRAN polypeptides can be
XX expressed by standard recombinant methodology. The polypeptides and
XX polynucleotides are useful in diagnosing, preventing or treating diseases
XX or conditions associated with the decreased expression or overexpression
XX of NTRAN, such as autoimmune/inflammatory, cardiovascular, neurological,
XX developmental, cell proliferative, transport, psychiatric, metabolic or
XX endocrine disorders. These diseases may include AIDS, allergies, atopic
XX dermatitis, arthritis, thyroiditis, obesity, Parkinson's disease,
XX Alzheimer's disease, stroke, epilepsy, myocardial infarction,
XX hypertension, cancer, atherosclerosis or hepatitis. There are also useful
XX in assessing the effects of exogenous compounds on the expression of
XX nucleic acid and amino acid sequences of NTRAN. The NTRAN or its
XX fragments are useful in screening compounds for effectiveness as agonist
XX or antagonist of the polypeptides, or in altering the expression of the
XX target polynucleotide and compounds that specifically bind to or modulate
XX the activity of the polypeptide. The microarray is useful in monitoring
XX or measuring protein-protein interactions, drug-target interactions, and
XX gene expression profiles. The present sequence represents a specific
XX example of a human NTRAN polypeptide.
XX
XX Sequence 301 AA;
XX

[illegible]

PR 06-DEC-2002; 2002US-0431455P.
 XX (INCY-) INCYTE CORP.
 XX
 XX Margula JP, Chawla NK, Beecha SD, Richardson TW, Hafalia AJA;
 PI Bhattacha UG, Burrill JD, Lee S, Blake JT, Ho A, Zheng W, Khare R;
 PI Ryan UK, Tang YT, Batra S, Burford N, Warren BA, Griffin JA;
 PI Ring HZ, Ison CH, Rillioet VS;
 XX
 DR WPI; 2004-294799/27.
 DR N-PSDB; ADL70546.
 XX
 PT New isolated G-protein coupled receptors (GCRBC) polypeptide and
 PT polynucleotide, useful for diagnosing and/or treating disorders with
 PT aberrant expression of GCRBC, such as arteriosclerosis, cirrhosis,
 PT cancer, stroke and AIDS.
 XX
 PS Claim 65; SEQ ID NO 4; 152pp; English.
 XX
 CC The present sequence is that of a Incyte Polypeptide ID 7519495CD1, i.e.
 CC human G-protein coupled receptor GCRBC-4. The polypeptide has homology to
 CC human and rat neurokinin A receptor. The invention provides GCRBC-1 to -
 CC 13 polypeptides ADL70530-ADL70542 and polynucleotides ADL70543-ADL70555,
 CC and method for using the polypeptides and polynucleotides for the
 CC detection, diagnosis and treatment of diseases, especially cell
 CC proliferation, neurological, cardiovascular, gastrointestinal,
 CC autoimmune/inflammatory, and metabolic disorders, and viral infections.
 CC They are also useful for facilitating the drug discovery process,
 CC including determination of efficacy, dosage, toxicity and pharmacology,
 CC and for investigating the pathogenesis of diseases and medical
 CC conditions. The receptors are also useful for identifying molecules
 CC involved in modulating taste or olfactory sensation.
 XX
 SQ Sequence 333 AA;
 XX
 Query Match 100.0%; Score 119; DB 8; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4.2e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYPIVNLALADLCMAAFNAFNP 23
 DB 69 NYPIVNLALADLCMAAFNAFNP 91
 XX
 RESULT 7
 ABM83530
 ID ABM83530 standard; protein; 333 AA.
 XX
 AC ABM83530;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Human diagnostic and therapeutic pproteins SEQ ID NO:3779.
 XX
 KM gene therapy; human diagnostic and therapeutic polynucleotide; dthp.
 XX
 OS Homo sapiens.
 XX
 PN WO2004023973-A2.
 XX
 PD 25-MAR-2004.
 XX
 PE 12-SEP-2003; 2003WO-US028227.
 XX
 PR 12-SEP-2002; 2002US-0410259P.
 PR 12-SEP-2002; 2002US-0410260P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 PI Hartschornme TA, Suchorolski MT, Altus CM, Pille SJ, Elder LV;
 PI Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;
 PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;

PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vilt UA, Kirton ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Patury S, Shi X, Suarez C;
 XX
 DR WPI; 2004-329368/30.
 DR N-PSDB; ACN42182.
 XX
 PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human
 PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 PS Claim 27; Page; 190pp; English.
 XX
 CC The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders, endocrine
 CC autoimmune/inflammatory disorder, developmental disorder, or
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dthp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dthp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 SQ Sequence 333 AA;
 XX
 Query Match 100.0%; Score 119; DB 8; Length 333;
 Best Local Similarity 100.0%; Pred. No. 4.2e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYPIVNLALADLCMAAFNAFNP 23
 DB 69 NYPIVNLALADLCMAAFNAFNP 91
 XX
 RESULT 8
 AAP90551
 ID AAP90551 standard; protein; 385 AA.
 XX
 AC AAP90551;
 XX
 DT 18-JUN-2003 (revised)
 DT 15-FEB-1990 (first entry)
 XX
 DE Bovine substance K receptor.
 XX
 KM Bovine substance K receptor; 5HT1c receptor.
 XX
 OS Bovidae.
 XX
 PN Location/Qualifiers
 XX
 FH 30..58
 FT /label= I
 FT /note= "transmembrane domain."
 FT 70..90
 FT /label= II
 FT /note= "transmembrane domain."
 FT 107..129
 FT /label= III
 FT /note= "transmembrane domain."
 FT 148..168
 FT /label= IV
 FT /note= "transmembrane domain"
 FT 194..218
 FT /label= V
 FT Domain

```
FT      Domain                               /note= "transmembrane domain."
PF      230
FT      /note= "cytoplasmic loop."
FT      252..274
FT      Domain                               /label= VI
FT      /note= "transmembrane domain."
FT      292..311
FT      Domain                               /label= VII
FT      /note= "transmembrane domain."
XX      MO6908149-A.
XX      08-SEP-1989.
XX      28-FEB-1989; 89WO-US000808.
XX      29-FEB-1988; 88US-00162654.
XX      18-JAN-1989; 89US-00298639.
XX      (UNCO ) UNIV COLUMBIA NEW YORK.
XX      Axel R, Jessell TM;
XX      WPI, 1989-278308/38.
XX      DNA encoding serotonin 5HT1c receptor - used for producing protein,
XX      antibodies and probes for studying receptor binding and screening drugs.
XX      Disclosure; Fig 11; 84pp; English.
XX      Sequence codes for bovine substance K receptor. See also AAP90549-P90554,
XX      AAP2111 and AAN90955. (Revised entry submitted on 18-JUN-2003 to correct
XX      the PN field.)
XX      Sequence 385 AA;
SQ      Query Match 100.0%; Score 119; DB 1; Length 385;
      Best Local Similarity 100.0%; Pred. No. 5e-11;
      Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 NYFIVNLALADLCMAAFNAAFNPF 23
      |||
      69 NYFIVNLALADLCMAAFNAAFNPF 91
Db
RESULT 9
ABU61827
ID      ABU61827 standard; protein; 387 AA.
XX      AC      ABU61827;
XX      DT      13-AUG-2003 (first entry)
XX      DE      Human substance K receptor.
XX      KW      Human; receptor; G-protein coupled receptor; inflammation; cancer;
XX      KW      substance K receptor; cell proliferation; metastasis; injury; tumour;
XX      KW      cell differentiation; heart function; nervous system development;
XX      KW      nervous system response.
XX      OS      Homo sapiens.
XX      FH      Key Location/Qualifiers
XX      FT      Misc-difference 231..243
XX      FT      /label= Unknown
XX      FT      /note= "These residues are omitted from the sequence"
XX      FT      Misc-difference 325..387
XX      FT      /label= Unknown
XX      FT      /note= "These residues are omitted from the sequence"
XX      PN      U86518414-B1.
XX      PD      11-FEB-2003.
```

```
XX      04-JAN-1999; 99US-00225024.
XX      15-FEB-1994; 94US-00196989.
XX      06-DEC-1996; 96US-00760936.
XX      (MACL/) MACLENNAN A J.
XX      MacLennan AJ;
XX      WPI, 2003-465606/44.
XX      New cDNA encodes a G-protein coupled receptor (designated pH218), useful
XX      for regulating cell proliferation and/or differentiation in vivo,
XX      particularly for treating e.g. cancers, or disorders related to heart
XX      function.
XX      Disclosure; Fig 2; 36pp; English.
XX      The invention relates to a new isolated polynucleotide molecule that
XX      comprises a polynucleotide that encodes a human G-protein coupled
XX      receptor (designated pH218). pH218 is a sphingosine-1-phosphate (S1P)
XX      receptor. The H218 polynucleotide or pH218 polypeptide is useful for
XX      regulating cell proliferation and/or differentiation, particularly in
XX      vivo. In particular, the H218 polynucleotide, or its encoded pH218
XX      polypeptide, is useful for treating cancers. The polynucleotide or
XX      polypeptide is also useful for treating disorders related to heart
XX      function, nervous system development, or nervous system responses to
XX      injury, tumorigenesis, metastases or inflammations. The pH218
XX      polypeptide functions as a growth factor ligand receptor. The present
XX      sequence represents the amino acid sequence of a human substance K
XX      receptor used to show homology with pH218 (see ACA62453)
XX      Sequence 387 AA;
SQ      Query Match 100.0%; Score 119; DB 7; Length 387;
      Best Local Similarity 100.0%; Pred. No. 5e-11;
      Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 NYFIVNLALADLCMAAFNAAFNPF 23
      |||
      69 NYFIVNLALADLCMAAFNAAFNPF 91
Db
RESULT 10
AAW80135
ID      AAW80135 standard; protein; 398 AA.
XX      AC      AAW80135;
XX      DT      17-DEC-1998 (first entry)
XX      DE      Human recombinant neurokinin-2 (NK-2) receptor protein.
XX      KW      Human; neurokinin receptor; NK-2; neurokinin A neurotransmitter;
XX      KW      abnormal smooth muscle cell contraction; asthma;
XX      KW      gastrointestinal disorder; peptic ulcer; ulcerative colitis.
XX      OS      Homo sapiens.
XX      FH      Key Location/Qualifiers
XX      FT      Protein 1..130
XX      FT      /note= "encoded by exon given in AAV66081"
XX      FT      1..31
XX      FT      /note= "extracellular domain"
XX      FT      60..67
XX      FT      /note= "intracellular domain"
XX      FT      90..103
XX      FT      /note= "extracellular domain"
XX      FT      130..149
XX      FT      /note= "intracellular domain"
XX      FT      /note= "intracellular domain"
XX      FT      131..196
XX      FT      /note= "encoded by exon given in AAV66082"
```


KW	Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
KM	constitutively activated GPCR; agonist; disease.
OS	Homo sapiens.
XX	Synthetic.
FN	M0200177172-A2.
PD	18-OCT-2001.
XX	
PF	05-APR-2001; 2001WO-US011098.
PR	
XX	07-APR-2000; 2000US-0195747P.
PA	(AREN-) ARENA PHARM INC.
XX	
PI	Lehmann-Bruntsma K, Liaw CM, Lin I;
DR	WPI; 2001-648759/74.
DR	N-PSDB; ABI98003.
XX	
PT	Identifying agonists of G protein-coupled receptors (GPCRs) for use in disease treatment, comprises contacting candidate compounds with versions of GPCRs.
XX	
PS	Claim 1; Page 328-329; 394pp; English.
XX	
CC	The invention relates to G protein-coupled receptors (GPCRs) for which the endogenous ligand has been identified. Non-endogenous constitutively activated versions of known GPCRs are used in the invention for the inverse identification of candidate compounds as receptor agonists, inverse agonists or partial agonists. Such agonists are useful as therapeutic agents for diseases or disorders associated with GPCRs. The present sequence is a non-endogenous version of a known human GPCR.
CC	
CC	
XX	
SO	Sequence 398 AA:
QY	
DB	
Query Match	100.0%; Score 119; DB 4; Length 398;
Best Local Similarity	100.0%; Pred. NO. 5.1e-11;
Matches 23; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
1 NYFVNLALADLCMAAFNAAFPNF 23	
69 NYFVNLALADLCMAAFNAAFPNF 91	
RESULT 12	
ABG32383	
ID	ABG32383 standard; protein; 398 AA.
XX	
AC	ABG32383;
XX	
DT	15-NOV-2002 (first entry)
XX	
XX	
XX	Human tachykinin receptor, TACR2, variant.
XX	
KW	Human; receptor; single nucleotide polymorphism; SNP; TACR2; tachykinin receptor.
XX	
OS	Homo sapiens.
XX	
XX	
FT	Key Location/Qualifiers
FT	Misc-difference 5 /note= "Asp is replaced by Gly as a result of a single
FT	nucleotide polymorphism"
FT	Misc-difference 23 /note= "Ile is replaced by Thr as a result of a single
FT	nucleotide polymorphism"
FT	Misc-difference 47 /note= "Ala is replaced by Thr as a result of a single
FT	nucleotide polymorphism"
FT	Misc-difference 251 /note= "Thr is replaced by Ala as a result of a single

Pt	Misc-difference	373	nucleotide polymorphism"
Ft	/note= "Thr is replaced by Ala as a result of a single nucleotide polymorphism"		
Ft	Misc-difference	385	/note= "Arg is replaced by His as a result of a single nucleotide polymorphism"
Ft	Misc-difference	395	/note= "His is replaced by Arg as a result of a single nucleotide polymorphism"
Pn	WO20063046-AI.		
Pd	15- AUG -2002 .		
Pf	09 -NOV -2001 ; 2001WO-US047394 .		
Pr	09 -NOV -2000 ; 2000US -0247649P .		
Xx	(GENA-) GENALISSANCE PHARM INC.		
Pa	Cappola G , Chew A , Gilson CR , Koshy B ; MPI , 2002 -636600 /68 .		
Dr	New genetic variants having polymorphisms in the Tachykinin receptor (TACR2) protein, useful for studying the function of TACR2, and for treating disorders associated with abnormal expression or function of TACR2 isogene.		
Pt	Claim 27; Page; 13pp; English.		
Pf	The invention relates to an isolated polypeptide comprising a polymeric variant of a reference sequence for the Tachykinin receptor (TACR2) protein. Also described is a method for: (1) haplotyping or genotyping the TACR2 gene of an individual; (2) predicting a haplotype pair for the TACR2 gene of an individual; (3) identifying an association between a trait and at least one haplotype or haplotype pair of the TACR2 gene; and (4) isolated oligonucleotide for detecting a single nucleotide polymorphism in the TACR2 gene. Polymorphic variants of the TACR2 gene are useful in studying the expression and biological function of TACR2, and in identifying drugs targeting TACR2 protein for treating disorders associated with abnormal expression or function of TACR2, e.g. asthma or breast cancer. Polynucleotides comprising a polymorphic gene variant or benefit may be used for therapeutic purposes, where a patient could benefit from expression or increased expression of a particular TACR2 protein isoform, or an expression vector encoding the isoform may be administered to the patient. Haplotype information is useful in improving the efficiency and output of several steps in drug discovery and development process, including target validation, identifying lead compounds, and early phase clinical trials. Information on polymorphisms may be applied in studying biological functions of TACR2 as well as in identifying drugs targeting this protein for the treatment of disorders related to its abnormal expression or function. The present sequence represents the amino acid sequence of human TACR2 variant protein comprising all possible isoforms caused by single nucleotide polymorphisms in the TACR2 coding sequence (see ABK91323 and ABK91324)		
Sq	Sequence 398 AA;		
Ov	Query Match	100.0%; Score 119; DB 5; Length 398; Best Local Similarity 100.0%; Pred. No. 5.le-11; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	1 NPYFVNLTADLCMAAFNAAPNF 23 69 NYFIYNLTLADLCMAAFNAAPNF 91 		
Result_13	ABG32382 standard; protein; 398 AA.		
ID	ABG32382 standard; protein; 398 AA.		
KX			

AC	ABG32382;
XX	
DT	15-NOV-2002 (first entry)
XX	
DE	Human tachykinin receptor, TACR2.
XX	
KM	Human; receptor; single nucleotide polymorphism; SNP; TACR2;
XX	tachykinin receptor.
OS	Homo sapiens.
XX	
PN	WO200263046-A1.
XX	
PD	15-AUG-2002.
PF	
XX	09-NOV-2001; 2001WO-US047394.
PR	
XX	09-NOV-2000; 2000US-0247649P.
PA	(GENA-) GENAISANCE PHARM INC.
PI	
DR	Cappola G, Chew A, Gilson CR, Koshy B;
XX	WPI; 2002-636600/68.
N-PSDB; ABK91323, ABK91324.	
PT	New genetic variants having polymorphisms in the Tachykinin receptor
PT	(TACR2) protein, useful for studying the function of TACR2, and for
PT	treating disorders associated with abnormal expression or function of
PT	TACR2 isogene.
PS	Claim 27; Fig 3; 139pp; English.
XX	
XX	The invention relates to an isolated polypeptide comprising a polymeric
CC	variant of a reference sequence for the Tachykinin receptor (TACR2)
CC	protein. Also described is a method for: (1) haplotyping or genotyping
CC	the TACR2 gene of an individual; (2) predicting a haplotype pair for the
CC	TACR2 gene of an individual; (3) identifying an association between a
CC	trait and at least one haplotype or haplotype pair of the TACR2 gene; and
CC	(4) isolated oligonucleotide for detecting a single nucleotide
CC	polymorphism in the TACR2 gene. Polymorphic variants of the TACR2 gene
CC	are useful in studying the expression and biological function of TACR2,
CC	and in identifying drugs targeting TACR2 protein for treating disorders
CC	associated with abnormal expression or function of TACR2, e.g. asthma or
CC	breast cancer. Polynucleotides comprising a polymorphic gene variant or
CC	bisectant may be used for therapeutic purposes, where a patient could
CC	benefit from expression or increased expression of a particular TACR2
CC	protein isoform, or an expression vector encoding the isoform may be
CC	administered to the patient. Haplotype information is useful in improving
CC	the efficiency and output of several steps in drug discovery and
CC	development process, including target validation, identifying lead
CC	compounds, and early phase clinical trials. Information on polymorphisms
CC	may be applied in studying biological functions of TACR2 as well as in
CC	identifying drugs targeting this protein for the treatment of disorders
CC	related to its abnormal expression or function. The present sequence
CC	represents the amino acid sequence of human TACR2
XX	
SQ	Sequence 398 AA;
Query Match	100.0%; Score 119; DB 5; Length 398;
Best Local Similarity	100.0%; Pred. No. 5,1e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	1 NYFIVNLALADLCMAAFNAATNF 23 NYFIVNLALADLCMAAFNAATNF 23
Db	69 NYFIVNLALADLCMAAFNAATNF 91
RESULT 14	
ID	ABP81985 standard; protein; 398 AA.
XX	
AC	ABP81985;

XX 04-MAR-2003 (first entry)
 DT Human tachykinin receptor 2 protein SEQ ID NO:456.
 XX
 XX
 XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KM G protein-coupled receptor modulator; antibody; immune-related disease;
 KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KM immunological-related disease; cell proliferative disease; autoimmune disease;
 KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KM osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KM ulcer.
 XX
 XX Homo sapiens.
 OS
 XX WO200261087-A2.
 PN
 XX 08-AUG-2002.
 PD
 XX 19-DEC-2001; 2001WO-US050107.
 PF
 XX 19-DEC-2000; 2000US-0257144P.
 PR
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 XX Burner GC, Roush CL, Brown JP;
 PI
 XX WPI; 2003-046718/04.
 DR
 XX N-PSDB; AB242833.
 DT
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 PT
 XX
 PS Disclosure; Fig 1; 523pp; English.
 XX
 XX The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, memory
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 CC
 XX
 SQ Sequence 398 AA;

Query Match 100.0%; Score 119; DB 6; Length 398;
 Best Local Similarity 100.0%; Pred. No. 5, 1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYEIVNALADLCMAAFNAFNF 23
 |||
 DB 69 NYEIVNALADLCMAAFNAFNF 91
 |||
 RESULT 15
 ID ADC86195 standard; protein; 398 AA.
 ADC86195
 XX
 XX ADC86195;
 AC
 XX
 XX 01-JAN-2004 (first entry)
 DT
 XX
 XX Human GPCR protein SEQ ID NO:648.
 DE
 XX
 XX human; GPCR; guanosine triphosphate-binding protein coupled receptor;
 KM gene therapy.
 KM
 XX
 XX Homo sapiens.
 OS
 XX EP1270724-A2.
 PN
 XX 02-JAN-2003.
 PD
 XX 18-JUN-2002; 2002EP-00013517.
 PF
 XX 18-JUN-2001; 2001JP-00246789.
 PR
 XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 PI
 XX Suwa M, Asai K, Akiyama Y, Aburatani H;
 PT
 XX WPI; 2003-315783/31.
 DR
 XX N-PSDB; ADC86194.
 DT
 XX New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 PT
 XX
 XX Claim 2; SEQ ID NO 648; 28pp; English.
 XX
 XX The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.
 CC
 XX
 SQ Sequence 398 AA;

Query Match 100.0%; Score 119; DB 7; Length 398;
 Best Local Similarity 100.0%; Pred. No. 5, 1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: January 23, 2006, 09:20:55
 Job time : 71.2 secs

This Page Blank (unused)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds
(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-9

Perfect score: 119

Sequence: 1 NYFIVNALADLCMAAFNAFNF 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	384	1 S00516	neurokinin 2 recep
2	119	100.0	398	1 J01059	neurokinin 2 recep
3	119	100.0	402	2 I56595	neurokinin 2 recep
4	114	95.8	384	2 I57957	neurokinin 2 recep
5	114	95.8	384	2 S20303	neurokinin 2 recep
6	114	95.8	390	2 A36737	neurokinin 2 recep
7	89	74.8	407	2 S20304	neurokinin 1 recep
8	89	74.8	407	2 A34357	neurokinin 1 recep
9	79	66.4	407	1 J01274	neurokinin 1 recep
10	79	66.4	407	2 S23510	neurokinin 1 recep
11	79	66.4	465	1 J01517	neurokinin 3 recep
12	79	66.4	504	2 A41783	neurokinin 3 recep
13	76	63.9	452	2 A34916	neurokinin 3 recep
14	75	63.0	385	2 S55524	neurokinin 3 recep
15	72	60.5	440	2 A44081	kappa-type opioid
16	72	60.5	419	2 S17783	neurokinin 3 recep
17	64	53.8	355	2 A42347	neurokinin 3 recep
18	64	53.8	355	2 I51319	neurokinin 3 recep
19	62	52.1	391	2 J02517	neurokinin 3 recep
20	62	52.1	501	2 J04447	neurokinin 3 recep
21	62	52.1	506	2 J07952	neurokinin 3 recep
22	62	52.1	560	2 A38731	neurokinin 3 recep
23	62	52.1	572	2 I39369	neurokinin 3 recep
24	61	51.3	374	2 S28285	neurokinin 3 recep
25	61	51.3	517	2 A45121	neurokinin 3 recep
26	59	49.6	484	2 S58868	neurokinin 3 recep
27	59	49.6	515	2 A40491	neurokinin 3 recep
28	59	49.6	515	2 J01525	neurokinin 3 recep
29	58	48.7	425	2 T15959	neurokinin 3 recep

30	58	48.7	601	2 JH0170
31	58	48.7	601	2 S12004
32	57	47.9	477	2 S71323
33	57	47.9	564	2 A38271
34	56	47.1	351	2 B45229
35	56	47.1	376	2 A48197
36	56	47.1	376	2 B48197
37	56	47.1	379	2 J06178
38	56	47.1	387	2 S55550
39	56	47.1	406	2 S55549
40	55	46.2	455	1 O00CG
41	54	45.4	345	2 T24659
42	54	45.4	349	1 J05490
43	54	45.4	387	2 I49246
44	54	45.4	429	2 S65656
45	54	45.4	444	2 C53886

ALIGNMENTS

RESULT 1
S00516
neurokinin 2 receptor - bovine
N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
C:Species: Bos primigenius taurus (cattle)
C>Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text change 09-Jul-2004
C/Accession: S00516
R:Masu, Y.; Nakayama, K.; Tamaki, H.; Harada, Y.; Kuno, M.; Nakanishi, S.
Nature 329, 836-838, 1997
A>Title: cDNA cloning of bovine substance K receptor through oocyte expression system.
A:Reference number: S00516; PMID:88039072; PMID:2823146
A:Accession: S00516
A:Molecule type: mRNA
A:Residues: 1-384 <MAS>
A:Cross-references: UNIPROT:P05363; UNIPARC:UP1000013018; EMBL:X06295; NID:9746; PIDN:
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane p
F/33-59/Domain: transmembrane #status predicted <TM1>
F/70-93/Domain: transmembrane #status predicted <TM2>
F/111-129/Domain: transmembrane #status predicted <TM3>
F/150-169/Domain: transmembrane #status predicted <TM4>
F/193-222/Domain: transmembrane #status predicted <TM5>
F/252-275/Domain: transmembrane #status predicted <TM6>
F/288-310/Domain: transmembrane #status predicted <TM7>
F/11,19/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/106-181/Dissulfide bonds: #status predicted

Query Match 100.0%; Score 119; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 3.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYFIVNALADLCMAAFNAFNF 23
DB 69 NYFIVNALADLCMAAFNAFNF 91

RESULT 2
J01059
neurokinin 2 receptor - human
N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
C:Species: Homo sapiens (man)
C/Date: 20-Apr-2000 #sequence, revision 20-Apr-2000 #text change 09-Jul-2004
C/Accession: J01059; J05075; A23658; A61224
R:Graham, A.; Hopkins, B.; Powell, S.J.; Danks, P.; Br1999, I.
Biochem. Biophys. Res. Commun. 177, 8-16, 1991
A>Title: Isolation and characterization of the human lung NK-2 receptor gene using rap1
A:Reference number: J01059; PMID:9125341; PMID:1170456
A:Accession: J01059
A:Molecule type: DNA
A:Residues: 1-398 <GRA1>
A:Cross-references: UNIPROT:P21452; UNIPARC:UP1000016ADE7; GB:M75105; NID:9189219; PIDN:
A:Accession: J05075

```

A:Molecule type: mRNA
A:Residues: 1-398 <GRA2>
A:Cross-references: UNIPARC:UPI000016ADE7; GB:M75105; NID:G189219; PIDN:AA805897.1; PID:
A:Experimental source: lung
A>Note: 23-ile and 375 His were also found
R:Gerard, N.P.; Eddy Jr., R.L.; Shows, T.B.; Gerard, C.
J. Biol. Chem. 265, 20455-20462, 1990
A>Title: The human neurokinin A (substance K) receptor. Molecular cloning of the gene, c
A:Reference number: A23658; MUID:91056095; PMID:2173708
A:Accession: A23658
A:Molecule type: DNA
A:Residues: 1-22, 'T', 24-240, 'L', 242-398 <GER>
A:Cross-references: UNIPARC:UPI000017351F; GB:M60284; GB:J05680; NID:G189140
A>Note: The authors translated the codon GGA for residue 317 as Glu
R:Kris, R.M.; South, V.; Saltzman, A.; Felder, S.; Rice, G.A.; Jaffe, M.; Huebner, K.; K
Cell Growth Differ. 2, 15-22, 1991
A>Title: Cloning and expression of the human substance K receptor and analysis of its ro
A:Reference number: A61224; MUID:91175483; PMID:1848773
A:Accession: A61224
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-22, 'T', 24-293, 'F', 294-374, 'H', 376-398 <KRI>
A:Cross-references: UNIPARC:UPI0000173520
C:Comment: The endogenous ligand of this receptor is neurokinin 2 (substance K or neurok
C:Genetics:
A:Gene: GDB:TAC2R
A:Cross-references: GDB:126367; OMIM:162321
A:Map position: 10q11-10q21
A:Introns: 131/2; 196/2; 247/3; 313/2
C:Superfamily: neurokinin 1 receptor
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane pr
F:33-59/Domain: transmembrane #status predicted <TM1>
F:70-93/Domain: transmembrane #status predicted <TM2>
F:111-122/Domain: transmembrane #status predicted <TM3>
F:150-169/Domain: transmembrane #status predicted <TM4>
F:194-222/Domain: transmembrane #status predicted <TM5>
F:252-275/Domain: transmembrane #status predicted <TM6>
F:288-310/Domain: transmembrane #status predicted <TM7>
F:11.19/finding site: carbohydrate (Asn) (covalent) #status predicted
F:106-181/disulfide bonds: #status predicted

Query Match          100.0%; Score 119; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 3.7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIWTLALADLCMAAFNAFNF 23
    |||||
Db 69 NYFIWTLALADLCMAAFNAFNF 91

RESULT 3
156595
neurokinin 2 receptor - guinea pig
C:Species: Cavia porcellus (guinea pig)
C>Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C:Accession: I56595
R:Baron, D.; Little, J.; Thomas, C.; Powell, S.; Downey-Jones, M.; Graham, A.
U. Recept. Res. 14, 399-421, 1994
A>Title: Isolation and characterization of neurokinin A receptor cDNAs from guinea-pig 1
A:Reference number: I56595; MUID:95182423; PMID:7877137
A:Accession: I56595
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-402 <RES>
A:Cross-references: UNIPROT:Q64077; UNIPARC:UPI00001301BF; GB:S76253; NID:G913374; PIDN:
C:Superfamily: neurokinin 1 receptor

Query Match          100.0%; Score 119; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIWTLALADLCMAAFNAFNF 23
    |||||

```

```

Db 69 NYFIWTLALADLCMAAFNAFNF 91

RESULT 4
157957
neurokinin 2 receptor - hamster
N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
C:Species: Cricetinae gen. sp. (hamster)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 19-May-2000
C:Accession: I57957
R:Baron, D.; Little, J.; Thomas, C.; Powell, S.; Berry, D.; Graham, A.
Mol. Pharmacol. 45, 9-19, 1994
A>Title: Isolation and pharmacological characterization of a hamster urinary bladder neu
A:Reference number: I57957; MUID:94134065; PMID:8302285
A:Accession: I57957
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-384 <RES>
A:Cross-references: UNIPARC:UPI00001301F1; GB:S68899; NID:G545230; PIDN:AA60680.1; PID:
C:Superfamily: neurokinin 1 receptor

Query Match          95.8%; Score 114; DB 2; Length 384;
Best Local Similarity 91.3%; Pred. No. 2.2e-10;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYFIWTLALADLCMAAFNAFNF 23
    |||||
Db 69 NYFIWTLALADLCMAAFNAFNF 91

RESULT 5
520303
neurokinin 2 receptor - mouse
N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
C:Species: Mus musculus (house mouse)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S20303
R:Sundelin, J.B.; Provvedini, D.M.; Wahlstedt, C.R.; Laurell, H.; Pohl, J.S.; Peterson,
Eur. J. Biochem. 203, 625-631, 1992
A>Title: Molecular cloning of the murine substance K and substance P receptor genes.
A:Reference number: S20303; MUID:92137253; PMID:1370937
A:Accession: S20303
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <SUN>
A:Cross-references: UNIPROT:P30549; UNIPARC:UPI0000029651; GB:X62933; NID:G54204; PIDN:
C:Superfamily: neurokinin 1 receptor
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match          95.8%; Score 114; DB 2; Length 384;
Best Local Similarity 91.3%; Pred. No. 2.2e-10;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYFIWTLALADLCMAAFNAFNF 23
    |||||
Db 69 NYFIWTLALADLCMAAFNAFNF 91

RESULT 6
A36737
neurokinin 2 receptor - rat
N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004
C:Accession: A36737
R:Isaai, Y.; Nakamishi, S.
Biochem. Biophys. Res. Commun. 165, 695-702, 1989
A>Title: Molecular characterization of rat substance K receptor and its mRNAs.
A:Reference number: A36737; MUID:90088481; PMID:2480781
A:Accession: A36737
A:Molecule type: mRNA
A:Residues: 1-390 <SAS>
A:Cross-references: UNIPROT:P16610; UNIPARC:UPI00001301F3; GB:M31838; NID:G206986; PIDN:

```

C/Superfamily: neurokinin 1 receptor
C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 95.8%; Score 114; DB 2; Length 390;
Best Local Similarity 91.3%; Pred. No. 2, 3e-10;

Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYPIVNLADLCMAAFNAFNF 23
|||:|||||:|||||
69 NYPIVNLADLCMAAFNAFNF 91

RESULT 7

neurokinin 1 receptor - mouse
N/Alternate names: NK-1 receptor; substance P receptor
C/Species: Mus musculus (house mouse)

C/Date: 22-Nov-1993 #sequence revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S20304; I56216; I73044

R/Sundell, J.B.; Provvedini, D.M.; Wahlestedt, C.R.; Laurell, H.; Pohl, J.S.; Peterson, Eur. J. Biochem. 203, 625-631, 1992

A/Title: Molecular cloning of the murine substance K and substance P receptor genes.
A/Reference number: S20303; MUID:92137253; PMID:1370937

A/Accession: S20304
A/Status: preliminary
A/Molecule type: mRNA

A/Residues: 1-407 <STN>
A/Cross-references: UNIPROT:P30548; UNIPARC:UPI000029850; GB:X62934; NID:954206; PIDN:C

R/Coak, G.A.; Elliott, D.; Metcalf, A.; Blum, A.M.; Sander, M.; Lynch, R.; Weinstein, J. J. Immunol. 152, 1830-1835, 1994

A/Title: Molecular evidence that granuloma T lymphocytes in murine schistosomiasis maned
A/Reference number: I56216; MUID:94165478; PMID:8120392

A/Accession: I56216
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Residues: 63-230 <COO>
A/Cross-references: UNIPARC:UPI000016CF5F; GB:I27826; NID:9450286; PIDN:AAAI7891.1; PID:

A/Experimental source: tissue brain
A/Accession: I73044

A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA

A/Residues: 63-230 <COO>
A/Cross-references: UNIPARC:UPI000016CF5F; GB:I27828; NID:9450290; PIDN:AAAI7892.1; PID:

A/Experimental source: tissue granuloma
A/Superfamily: neurokinin 1 receptor

C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 74.8%; Score 89; DB 2; Length 407;
Best Local Similarity 69.6%; Pred. No. 2, 2e-06;

Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYPIVNLADLCMAAFNAFNF 23
|||:|||||:|||||
68 NYPIVNLADLCMAAFNAFNF 90

RESULT 8

neurokinin 1 receptor - rat
N/Alternate names: NK-1 receptor; substance P receptor
C/Species: Rattus norvegicus (Norway rat)

C/Date: 22-Jan-1993 #sequence revision 22-Jan-1993 #text_change 09-Jul-2004
C/Accession: A38692; A34357; A40089

R/Hershey, A.D.; Dykema, P.E.; Krause, J.B. J. Biol. Chem. 266, 4366-4374, 1991

A/Title: Organization, structure, and expression of the gene encoding the rat substance
A/Reference number: A38692; MUID:91154239; PMID:1705552

A/Accession: A38692
A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-407 <HR>
A/Cross-references: UNIPROT:P14600; UNIPARC:UPI00001301E4; GB:M34751

R/Yokota, Y.; Sasaki, Y.; Tanaka, K.; Fujiwara, T.; Tsuchida, K.; Shigemoto, R.; Kakizuka J. Biol. Chem. 264, 17649-17652, 1989

A/Title: Molecular characterization of a functional cDNA for rat substance P receptor.
A/Reference number: A34357; MUID:90036822; PMID:2478537

A/Accession: A34357
A/Molecule type: mRNA

A/Residues: 1-407 <YOK>
A/Cross-references: UNIPARC:UPI00001301E4; GB:J05097; NID:9207051; PIDN:AAA42175.1; PID

R/Hershey, A.D.; Krause, J.B. Science 247, 958-962, 1990

A/Title: Molecular characterization of a functional cDNA encoding the rat substance P r:
A/Reference number: A40089; MUID:90161991; PMID:2154852

A/Accession: A40089
A/Status: preliminary
A/Molecule type: mRNA

A/Residues: 1-74, 'R', '76-212', 'A', '214-407' <HR>
A/Cross-references: UNIPARC:UPI0000176611; GB:M31477

C/Superfamily: neurokinin 1 receptor
C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 74.8%; Score 89; DB 2; Length 407;
Best Local Similarity 69.6%; Pred. No. 2, 2e-06;

Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYPIVNLADLCMAAFNAFNF 23
|||:|||||:|||||
68 NYPIVNLADLCMAAFNAFNF 90

RESULT 9

neurokinin 1 receptor - human
N/Alternate names: NK-1 receptor; substance P receptor
C/Species: Homo sapiens (man)

C/Date: 20-Apr-2000 #sequence revision 20-Apr-2000 #text_change 09-Jul-2004
C/Accession: A41134; J01274; J04478; S21188

R/Gerard, N.P.; Garraway, L.A.; Bddy Jr., R.L.; Shows, T.B.; Iijima, H.; Paquet, J.L.; Biochemistry 30, 10640-10646, 1991

A/Title: Human substance P receptor (NK-1): organization of the gene, chromosome local:
A/Reference number: A41134; MUID:9201510; PMID:1657150

A/Accession: A41134
A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-328, 'G', '329-332', '334-407' <GHR>
A/Cross-references: UNIPROT:P25103; UNIPARC:UPI00001751B; GB:M76675; NID:9189231

A/Note: In the authors' translation 333-Gly is shown before residue 329 and, consequent
R/Takeda, Y.; Chou, K.B.; Takeda, J.; Sackas, B.S.; Krause, J.B. Biochem. Biophys. Res. Commun. 179, 1232-1240, 1991

A/Title: Molecular cloning, structural characterization and functional expression of th
A/Reference number: J01274; MUID:92028856; PMID:1718267

A/Accession: J01274
A/Status: preliminary
A/Molecule type: mRNA

A/Residues: 1-407 <TKK1>
A/Cross-references: UNIPARC:UPI000039D68; GB:M74290; NID:9338612; PIDN:AAA60601.1; PID

R/Hopkins, B.; Powell, S.J.; Danks, P.; Brigg, I.; Graham, A. Biochem. Biophys. Res. Commun. 180, 1110-1117, 1991

A/Title: Isolation and characterization of the human lung NK-1 receptor cDNA.
A/Reference number: J04478; MUID:92062052; PMID:1659396

A/Accession: J04478
A/Status: preliminary
A/Molecule type: mRNA

A/Residues: 1-407 <HOP>
A/Cross-references: UNIPARC:UPI000039D68; GB:S62045; NID:9237994; PIDN:AAAB20168.1; PID

A/Experimental source: lung
A/Note: The authors translated the codon CAA for residue 31 as Glu

R/Takahashi, K.; Tanaka, A.; Hara, M.; Nakashima, S. Eur. J. Biochem. 204, 1025-1033, 1992

A/Title: The primary structure and gene organization of human substance P and neuromedi:
A/Reference number: S21188; MUID:92201186; PMID:1312928

A/Accession: S21188
A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-407 <TKK2>
A/Cross-references: UNIPARC:UPI000039D68; GB:X65177; NID:936636; PIDN:CAA46292.1; PID:

C/Comment: The endogenous ligand of this receptor is neurokinin 1 (substance P), one of
C/Genetics:

A/Gene: GDB:TMCR

A:Cross-references: GDB:128977; OMIM:162323
 A:Map position: 2pter-2qter
 C:Superfamily: neurokinin 1 receptor
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane protein
 F:32-58/Domain: transmembrane #status predicted <TM1>
 F:69-92/Domain: transmembrane #status predicted <TM2>
 F:110-128/Domain: transmembrane #status predicted <TM3>
 F:149-168/Domain: transmembrane #status predicted <TM4>
 F:195-221/Domain: transmembrane #status predicted <TM5>
 F:249-273/Domain: transmembrane #status predicted <TM6>
 F:286-308/Domain: transmembrane #status predicted <TM7>
 F:14,18/binding site: carbohydrate (Asn) (covalent) #status predicted
 F:105-180/disulfide bonds: #status predicted

Query Match 66.4%; Score 79; DB 1; Length 407;
 Best Local Similarity 65.2%; Pred. No. 8.4e-05;
 Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 NYFVNLALADLCMAAFNAFNP 23
 |||:||||:|||||
 Db 68 NYFLVNLAFASMAAFNTVNF 90

RESULT 10
 S23510
 neurokinin 1 receptor - guinea pig
 N:Alternate names: NK-1 receptor; substance P receptor
 C:Species: Cavia porcellus (guinea pig)
 C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
 A:Accession: S23510; S19198
 R:Gorbulov, V.; Akhundova, A.; Iuzius, H.; Fahrenholz, F.
 Biochem. Biophys. Acta 1131, 99-102, 1992
 A:Title: Molecular cloning of substance P receptor cDNA from guinea-pig uterus.
 A:Reference number: S23510; MUID:92256498; PMID:1374648
 A:Accession: S23510
 A:Molecule type: mRNA
 A:Residues: 1-407 <GOR>
 A:Cross-references: UNIPROT:P30547; UNIPARC:UPI0000001112; EMBL:X64323; NID:G949565; PIDN:
 C:Superfamily: neurokinin 1 receptor
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:32-55/Domain: transmembrane #status predicted <TM1>
 F:69-89/Domain: transmembrane #status predicted <TM2>
 F:117-128/Domain: transmembrane #status predicted <TM3>
 F:149-169/Domain: transmembrane #status predicted <TM4>
 F:196-217/Domain: transmembrane #status predicted <TM5>
 F:250-280/Domain: transmembrane #status predicted <TM6>
 F:289-308/Domain: transmembrane #status predicted <TM7>

Query Match 66.4%; Score 79; DB 2; Length 407;
 Best Local Similarity 65.2%; Pred. No. 8.4e-05;
 Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 NYFVNLALADLCMAAFNAFNP 23
 |||:||||:|||||
 Db 68 NYFLVNLAFASMAAFNTVNF 90

RESULT 11
 JQ1517
 neurokinin 3 receptor - human
 N:Alternate names: neuromedin K receptor; NK-3 receptor
 C:Species: Homo sapiens (man)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
 A:Accession: JQ1517; S20435; S21237
 R:Huang, R.R.C.; Cheung, A.H.; Mazina, K.E.; Strader, C.D.; Fong, T.M.
 Biochem. Biophys. Res. Commun. 184, 966-972, 1992
 A:Title: cDNA sequence and heterologous expression of the human neurokinin-3 receptor.
 A:Reference number: JQ1517; MUID:92246993; PMID:1374246
 A:Accession: JQ1517
 A:Molecule type: mRNA
 A:Status: nucleic acid sequence not shown
 A:Residues: 1-465 <HUA>
 A:Cross-references: UNIPROT:P29371; UNIPARC:UPI0000050413; GB:M89473; NID:G189223; PIDN:

A:Experimental source: brain
 R:Bull, G.; Schultz, M.F.; Arkinsteal, S.J.; Maury, K.; Misossten, M.; Adami, N.; Talabo
 RBBS Lett. 299, 90-95, 1992
 A:Title: Molecular characterization, expression and localization of human neurokinin-3
 A:Reference number: S20435; MUID:92183914; PMID:1312036
 A:Accession: S20435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2, '1', 4-62, 'R', 64-465 <BUE>
 A:Cross-references: UNIPARC:UPI00001684D2; GB:S86397; NID:9246908; PIDN:AA821706.1; PID:
 R:Yakushiji, K.; Tanaka, A.; Hara, M.; Nakashima, S.
 Eur. J. Biochem. 204, 1025-1033, 1992
 A:Title: The primary structure and gene organization of human substance P and neuromedin
 A:Reference number: S21188; MUID:92201186; PMID:1312928
 A:Accession: S21237
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-438, 'P', 440-465 <TAK>
 A:Cross-references: UNIPARC:UPI00016ADCE; GB:X65172; NID:G35022; PIDN:CAA46291.1; PID:G
 C:Comment: The endogenous ligand of this receptor is neurokinin 3 (neuromedin K), one of
 C:Genetics:
 A:Gene: GDB:TACR3
 A:Cross-references: GDB:9599126
 C:Superfamily: neurokinin 1 receptor
 C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane protein
 F:83-111/Domain: transmembrane #status predicted <TM1>
 F:122-147/Domain: transmembrane #status predicted <TM2>
 F:160-181/Domain: transmembrane #status predicted <TM3>
 F:202-221/Domain: transmembrane #status predicted <TM4>
 F:247-272/Domain: transmembrane #status predicted <TM5>
 F:300-321/Domain: transmembrane #status predicted <TM6>
 F:333-355/Domain: transmembrane #status predicted <TM7>
 F:23, 50, 73/binding site: carbohydrate (Asn) (covalent) #status predicted
 F:158-233/disulfide bonds: #status predicted

Query Match 66.4%; Score 79; DB 1; Length 465;
 Best Local Similarity 65.2%; Pred. No. 9.5e-05;
 Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 NYFVNLALADLCMAAFNAFNP 23
 |||:||||:|||||
 Db 121 NYFLVNLAFSDASMAAFNTVNF 143

RESULT 12
 A41783
 tachykinin receptor NKD - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Apr-2000
 A:Accession: A41783
 R:Monnier, D.; Colas, J.F.; Rosay, P.; Hen, R.; Borrelli, E.; Maroteaux, L.
 J. Biol. Chem. 267, 1298-1302, 1992
 A:Title: NKD, a developmentally regulated tachykinin receptor in Drosophila.
 A:Reference number: A41783; MUID:92112776; PMID:1370464
 A:Accession: A41783
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-504 <MON>
 A:Cross-references: UNIPARC:UPI0000176634
 A:Note: sequence extracted from NCBI Backbone (NCBI:76545)
 C:Genetics:
 A:Gene: FlyBase:TAKR86C
 A:Cross-references: FlyBase:FBgn0004841
 C:Superfamily: neurokinin 1 receptor
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 66.4%; Score 79; DB 2; Length 504;
 Best Local Similarity 56.5%; Pred. No. 0.0001;
 Matches 13; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NYFVNLALADLCMAAFNAFNP 23
 |||:||||:|||||
 Db 119 NYFLNLISIDLMSLNCVFNF 141

RESULT 13

A34916

neurokinin 3 receptor - rat

M/Alternate names: neuromedin K receptor; NK-3 receptor

C/Species: Rattus norvegicus (Norway rat)

C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004

C/Accession: A34916

R/Shigemoto, R.; Yokota, Y.; Tsuchida, K.; Nakanishi, S.

J. Biol. Chem. 265, 623-628, 1990

A/Title: Cloning and expression of a rat neuromedin K receptor cDNA.

A/Reference number: A34916; MUID:90110113; PMID:2153106

A/Accession: A34916

A/Molecule type: mRNA

A/Residues: 1-452 <SH1>

A/Cross-references: UNIPROT:P16177; UNIPARC:UPI00001301F7; GB:J05189; NID:G205670; PIDN:

C/Superfamily: neurokinin 1 receptor

C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match

63.9%; Score 76; DB 2; Length 452;

Best Local Similarity 60.9%; Pred. No. 0.00028;

Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 NYFVNLALADLCMAAFNAFNP 23

DB 108 NYFVNLALAFSDASVAAPNTLVNF 130

RESULT 14

855524

neurokinin 3 receptor - mouse (fragment)

N/Alternate names: neuromedin K receptor; NK-3 receptor

C/Species: Mus musculus (house mouse)

C/Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004

C/Accession: 855524; I73045

R/Maroteaux, L.

submitted to the EMBL Data Library, June 1995

A/Reference number: 855524

A/Accession: 855524

A/Molecule type: mRNA

A/Residues: 1-385 <MAR>

A/Cross-references: UNIPROT:P47937; UNIPARC:UPI000002783F; EMBL:X87823; NID:G861055; PID

R/Cook, G.A.; Elliott, D.; Metwalli, A.; Blum, A.M.; Sandor, M.; Lynch, R.; Weinstock, J.

J. Immunol. 152, 1830-1835, 1994

A/Title: Molecular evidence that granuloma T lymphocytes in murine schistosomiasis manec

A/Reference number: 156216; MUID:94165478; PMID:8120392

A/Accession: I73045

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 103-197 'S', 199-266 'P', 268-328 <COO>

A/Cross-references: UNIPARC:UPI000016CF60; GB:L27827; NID:G450288; PIDN:AAA17893.1; PID:

C/Superfamily: neurokinin 1 receptor

Query Match

63.0%; Score 75; DB 2; Length 385;

Best Local Similarity 60.9%; Pred. No. 0.00034;

Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 NYFVNLALADLCMAAFNAFNP 23

DB 108 NYFVNLALAFSDASVAAPNTLVNF 130

RESULT 15

A44081

kappa-type opioid receptor - human

C/Species: Homo sapiens (man)

C/Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004

C/Accession: A44081

R/Xie, G.X.; Miyajima, A.; Goldstein, A.

Proc. Natl. Acad. Sci. U.S.A. 89, 4124-4128, 1992

A/Title: Expression cloning of cDNA encoding a seven-helix receptor from human placenta

A/Reference number: A44081; MUID:92237319; PMID:1315051

A/Accession: A44081

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-440 <XIE>

A/Cross-references: UNIPROT:P30098; UNIPARC:UPI000000F07; GB:M84605; NID:G189391; PIDN:

C/Superfamily: neurokinin 1 receptor

C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match

60.5%; Score 72; DB 2; Length 440;

Best Local Similarity 65.2%; Pred. No. 0.0012;

Matches 15; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 NYFVNLALADLCMAAFNAFNP 23

DB 96 NSFLVNLAFADAAALNLALVNF 118

Search completed: January 23, 2006, 09:32:53

Job time : 9.1 secs

This page blank (copy)

AC P05363; 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2 receptor)
DB (NK-2R) (Tachykinin receptor 2).
GN Name=TACR2; Synonyms=TACR;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86039072; PubMed=282316; DOI=10.1038/329836a0.
RA Masu Y., Nakayama K., Tamaki H., Harada Y., Kuno M., Nakanishi S.;
RT "cDNA cloning of bovine substance-K receptor through oocyte expression
system";
RL Nature 329:836-838 (1987).
CC -!- FUNCTION: This is a receptor for the tachykinin neuropeptide
substance K (neurokinin A). It is associated with G proteins that
activate a phosphatidylinositol-calcium second messenger system.
CC The rank order of affinity of this receptor to tachykinins is:
CC substance K > neuromedin K > substance P.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL: X06295; CA29621.1; -; mRNA.
DR PIR: S00516; S00516.
DR HSP: P02699; 1F88.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001681; Neurokinin_receptor.
DR InterPro: IPR000913; NK2_receptor.
DR PANTHER: PTHR19264:SF70; NK2_receptor; 1.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PRINTS: PR01025; NEUROKININ2R.
DR PRINTS: PR00244; NEUROKININR.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECPT_FL_2; 1.
KW G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
KW Phosphorylation; Receptor; Transducer; Transmembrane.
FT TOPO_DOM 1 32
FT TRANSMEM 33 56
FT TOPO_DOM 57 69
FT TRANSMEM 70 90
FT TOPO_DOM 91 107
FT TRANSMEM 108 129
FT TOPO_DOM 130 149
FT TRANSMEM 150 170
FT TOPO_DOM 171 196
FT TRANSMEM 197 218
FT TOPO_DOM 219 251
FT TRANSMEM 252 272
FT TOPO_DOM 273 290
FT TRANSMEM 291 310
FT TOPO_DOM 311 364
FT TRANSMEM 364 384
FT LIPID 324 324
FT CARBOHYD 11 11
FT CARBOHYD 19 19
FT CARBOHYD 18 18
FT DISULFID 106 181
SQ SEQUENCE 384 AA; 43069 MW; 1EC676BBD7FF1672 CRC64;
Query Match 100.0%; Score 119; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFVNLALADLCMAAFNAFNF 23
Db |||||
69 NYFVNLALADLCMAAFNAFNF 91
RESULT 3
ID NC2R CANFA STANDARD; PRT; 384 AA.
AC 05DUB2;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2 receptor)
DB (NK-2R) (Tachykinin receptor 2).
GN Name=TACR2;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_TaxID=9615;
[1]
RP NUCLEOTIDE SEQUENCE [mRNA].
RA Engberg S., Drmota F.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This is a receptor for the tachykinin neuropeptide
substance K (neurokinin A). It is associated with G proteins that
activate a phosphatidylinositol-calcium second messenger system
(by similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL: AJ884916; CA158657.1; -; mRNA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR001681; Neurokinin_receptor.
DR InterPro: IPR000913; NK2_receptor.
DR PANTHER: PTHR19264:SF70; NK2_receptor; 1.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PRINTS: PR01025; NEUROKININ2R.
DR PRINTS: PR00244; NEUROKININR.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECPT_FL_2; 1.
KW G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
KW Phosphorylation; Receptor; Transducer; Transmembrane.
FT TOPO_DOM 1 32
FT TRANSMEM 33 56
FT TOPO_DOM 57 69
FT TRANSMEM 70 90
FT TOPO_DOM 91 107
FT TRANSMEM 108 129
FT TOPO_DOM 130 149
FT TRANSMEM 150 170
FT TOPO_DOM 171 196
FT TRANSMEM 197 218
FT TOPO_DOM 219 251
FT TRANSMEM 252 272
FT TOPO_DOM 273 290
FT TRANSMEM 291 310
FT TOPO_DOM 311 384
FT TRANSMEM 384 384
FT LIPID 324 324
FT CARBOHYD 11 11
FT CARBOHYD 18 18
FT CARBOHYD 19 19
FT DISULFID 106 181
SQ SEQUENCE 384 AA; 43359 MW; A7C3D524ABEC2BDF CRC64;
Query Match 100.0%; Score 119; DB 1; Length 384;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYPIVNLADLCMAAFNAFNF 23
 |||||
 DB 69 NYPIVNLADLCMAAFNAFNF 91

RESULT 4
 NCNR HUMAN STANDARD; PRT; 398 AA.
 ID NK2R_HUMAN
 AC P21452; Q9UDE7; 18, Created
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2 receptor)
 DB (NK-2R) (Tachykinin receptor 2).
 GN Name: TACR2; Synonyms: NK2R, TAC2R;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Stomach, and Trachea;
 RX MEDLINE=91056095; PubMed=2173708;
 RA Gerard N.P., Eddy R.L., Jr., Shows T.B., Gerard C.;
 RT "The human neurokinin A (substance K) receptor. Molecular cloning of
 RT the gene, cDNA sequence, localization, and isolation of cDNA from tracheal
 RT and gastric tissues."
 RL J. Biol. Chem. 265:20455-20462 (1990).
 RN [2]
 RP TISSUE=Stomach.
 RX MEDLINE=91093253; PubMed=1845990;
 RA Gerard N.P., Eddy R.L., Jr., Shows T.B., Gerard C.;
 RL J. Biol. Chem. 266:1354-1354 (1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE, AND VARIANT THR-23.
 RC TISSUE=Lung;
 RX MEDLINE=91254341; PubMed=1710456;
 RA Graham A., Hopkins B., Powell S.J., Danks P., Briggs I.;
 RT "Isolation and characterization of the human lung NK-2 receptor gene
 RT using rapid amplification of cDNA ends."
 RL Biochem. Biophys. Res. Commun. 177:8-16 (1991).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92060616; PubMed=1659296;
 RA Gerard N.P., Gerard C.;
 RT "Molecular cloning of the human neurokinin-2 receptor cDNA by
 RT polymerase chain reaction and isolation of the gene."
 RL Ann. N. Y. Acad. Sci. 632:389-390 (1991).
 RN [5]
 RP NUCLEOTIDE SEQUENCE, AND VARIANT HIS-375.
 RX PubMed=1848773;
 RA Kiehl R.M., South V., Saltzman A., Felder S., Ricca G.A., Jaye M.,
 RA Heuber K., Kagan J., Croce C.M., Schlesinger J.;
 RT "Cloning and expression of the human substance K receptor and analysis
 RT of its role in mitogenesis."
 RL Cell Growth Differ. 2:15-22 (1991).
 RN [6]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA).
 RC TISSUE=Stomach;
 RA Kopar S.A., Aronstam R.S., Sharma S.V.;
 RT "cDNA clones of human proteins involved in signal transduction
 RT sequenced by the Gutthrie cDNA resource center (www.cdna.org).";
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
 CC substance K (neurokinin A). It is associated with G proteins that
 CC activate a phosphatidylinositol-calcium second messenger system.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- MISCELLANEOUS: The rank order of affinity of this receptor to
 CC tachykinins is: substance K > neurokinin K > substance P.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC EMBL, M57414; AAC31760.1; -; mRNA.
 CC EMBL, M60284; AAA60347.1; -; Genomic DNA.
 CC EMBL, M57415; AAA60347.1; JOINED; Genomic DNA.
 CC EMBL, M60281; AAA60347.1; JOINED; Genomic DNA.
 CC EMBL, M60282; AAA60347.1; JOINED; Genomic DNA.
 CC EMBL, M60283; AAA60347.1; JOINED; Genomic DNA.
 CC EMBL, M75105; AAB05897.1; -; Genomic DNA.
 CC EMBL, M75101; AAB05897.1; JOINED; Genomic DNA.
 CC EMBL, M75102; AAB05897.1; JOINED; Genomic DNA.
 CC EMBL, M75103; AAB05897.1; JOINED; Genomic DNA.
 CC EMBL, M75104; AAB05897.1; JOINED; Genomic DNA.
 CC EMBL, AY322545; AAB84358.1; -; mRNA.
 CC PIR, JQ1059; JQ1059.
 CC HSSP, P02699; 1P88.
 CC DR HGNCL; HGNCL:11527; TACR2.
 CC DR HGNCL; HGNCL:11527; TACR2.
 CC DR MIM; 163321; -.
 CC DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 CC DR GO; GO:0004995; P: tachykinin receptor activity; TAS.
 CC DR GO; GO:0007588; P: excretion; TAS.
 CC DR GO; GO:0006936; P: muscle contraction; TAS.
 CC DR GO; GO:0007217; P: tachykinin signaling pathway; TAS.
 CC DR InterPro; IPR000276; GPCR Rhodopsin.
 CC DR InterPro; IPR001681; Neurokinin receptor.
 CC DR InterPro; IPR000913; NK2_receptor.
 CC DR PANTHER; PTHR19264; SF70; NK2_receptor; 1.
 CC DR Pfam; PF00001; 7tm_1; 1.
 CC DR PRINTS; PR00237; GPCRHHODOPSIN.
 CC DR PRINTS; PR01025; NEUROKININR.
 CC DR PRINTS; PR00244; NEUROKININR.
 CC DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 CC DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
 CC KW G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
 CC KM Phosphorylation; Polymorphism; Receptor; Transducer; Transmembrane.
 CC FT TOPO_DOM 1 32
 CC FT TRANSMEM 33 56
 CC FT TOPO_DOM 57 69
 CC FT TRANSMEM 70 90
 CC FT TOPO_DOM 91 107
 CC FT TRANSMEM 108 129
 CC FT TOPO_DOM 130 149
 CC FT TRANSMEM 150 170
 CC FT TOPO_DOM 171 196
 CC FT TRANSMEM 197 218
 CC FT TOPO_DOM 219 251
 CC FT TRANSMEM 252 272
 CC FT TOPO_DOM 273 290
 CC FT TRANSMEM 291 310
 CC FT TOPO_DOM 311 338
 CC FT LIPID 324 324
 CC FT CARBOHYD 11 11
 CC FT CARBOHYD 19 19
 CC FT DISULFID 106 181
 CC FT VARIANT 23 23
 CC FT VARIANT 23 23
 CC FT VARIANT 375 375
 CC FT VARIANT 375 375
 CC FT CONFLICT 38 38
 CC FT CONFLICT 293 293
 CC FT SEQUENCE 398 AA; 44439 MW; 7AF15BB137463A1 CRC64;
 CC Query Match 100.0%; Score 119; DB 1; Length 398;
 CC Best Local Similarity 100.0%; Pred. No. 8-2e-10;
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NYPIVNLADLCMAAFNAFNF 23

```

Db      69 NYFIVNALADLCMAAFNAFNF 91
|||||
RESULT 5
Q8NG08 HUMAN PRELIMINARY; PRT; 398 AA.
ID Q8NG08
AC Q8NG08
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Seven transmembrane helix receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxId=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Teureunt S., Aburatani H., Asai K., Akiyama Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AB065731; BAC05952.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005886; C:plasma membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0004995; F:tachykinin receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001681; Neurokin receptor.
DR InterPro; IPR000913; NK2_receptor.
DR PANTHER; PTHR19264:SF70; NK2_receptor; 1.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR01025; NEUROKININR.
DR PRINTS; PR00244; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_REC_P1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_REC_P1_2; 1.
DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 398 AA; 44442 MW; 20AD618EF74263A7 CRC64;

Query Match 100.0%; Score 119; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 8.2e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIVNALADLCMAAFNAFNF 23
|||||
Db 69 NYFIVNALADLCMAAFNAFNF 91

```

```

RL Ann. N. Y. Acad. Sci. 632:426-427(1991).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR Ensembl; ENSG00000075073; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005886; C:plasma membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0004995; F:tachykinin receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001681; Neurokin receptor.
DR InterPro; IPR000913; NK2_receptor.
DR PANTHER; PTHR19264:SF70; NK2_receptor; 1.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR01025; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_REC_P1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_REC_P1_2; 1.
DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 398 AA; 44356 MW; 39C170237A6D6E7E CRC64;

Query Match 100.0%; Score 119; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 8.2e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIVNALADLCMAAFNAFNF 23
|||||
Db 69 NYFIVNALADLCMAAFNAFNF 91

RESULT 7
Q4QR15 HUMAN PRELIMINARY; PRT; 398 AA.
ID Q4QR15
AC Q4QR15
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Tachykinin receptor 2.
GN Name=TRACR2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxId=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE-G-protein coupled receptors;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Shamen C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshlyuki S., Carinici P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kerteman M., Madan A.C., Rodriques S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalske U., Smalies D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE-G-protein coupled receptors;

```

RG NIH MGC Project;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DR EMBL, BC096842; AAH96842.1; -; mRNA.
DR InterPro, IPR000276; GPCR_Rhodopsn.
DR InterPro, IPR001681; Neurokin_receptor.
DR InterPro, IPR000913; NK2_receptor.
DR Pfam, PF00001; 7tm.1; 1.
DR PRINTS, PR00237; GPCRHOOPS.
DR PRINTS, PR01025; NEUROKININR.
DR PRINTS, PR00244; NEUROKININR.
DR PRINTS, PR00237; G_PROTEIN_RECP_P1_1; 1.
DR PROSITE, PS50262; G_PROTEIN_RECP_P1_2; 1.
DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.
KW SEQUENCE 398 AA; 44482 MW; 619D59B32CA4A6AD CRC64;

Query Match 100.0%; Score 119; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 8.2e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFYVNLALADLCMAAFNAFNF 23
Db 69 NYFYVNLALADLCMAAFNAFNF 91

RESULT 8
NK2R_CAVPO STANDARD; PRT; 402 AA.
ID NK2R_CAVPO
AC 064077;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2 receptor)
DE (NK-2R) (Tachykinin receptor 2).
GN Name=TACR2;
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
OC Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxId=10141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Dunkin-Hartley; TISSUE=Lung;
RX MEDLINE=95182423; PubMed=7877137;
RA Aharony D., Little J., Thomas C., Powell S., Downey-Jones M.,
RA Graham A.;
RT "Isolation and characterization of neurokinin A receptor cDNAs from
RT guinea-pig lung and rabbit pulmonary artery."
RL J. Recept. Res. 14:399-421(1994).
CC -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
CC substance K (neurokinin A). It is associated with G proteins that
CC activate a phosphatidylinositol-calcium second messenger system.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL, S76253; AAB33553.1; -; mRNA.
CC PIR, I56595; I56595.
CC HSBP, P02699; I988.
CC InterPro, IPR000276; GPCR_Rhodopsn.
CC InterPro, IPR001681; Neurokin_receptor.
CC InterPro, IPR000913; NK2_receptor.
CC PANTHER, PTHR19264:SP70; NK2_receptor; 1.
CC Pfam, PF00001; 7tm.1; 1.
CC PRINTS, PR00237; GPCRHOOPS.
CC PRINTS, PR01025; NEUROKININR.
CC PRINTS, PR00244; NEUROKININR.

DR PROSITE, PS00237; G_PROTEIN_RECP_P1_1; 1.
DR PROSITE, PS50262; G_PROTEIN_RECP_P1_2; 1.
KW G-protein coupled receptor; Transducer; Transmembrane.
KW Phosphorylation; Receptor; Transducer; Transmembrane.
FT TOPO_DOM 1 32
FT TOPO_DOM 33 56
FT TOPO_DOM 57 69
FT TOPO_DOM 70 90
FT TOPO_DOM 91 107
FT TOPO_DOM 108 129
FT TOPO_DOM 130 149
FT TOPO_DOM 150 170
FT TOPO_DOM 171 196
FT TRANSMEM 197 218
FT TOPO_DOM 219 251
FT TOPO_DOM 252 272
FT TRANSMEM 273 290
FT TOPO_DOM 291 310
FT TOPO_DOM 311 402
FT LIPID 324 324
FT CARBOHYD 11 11
FT CARBOHYD 19 19
FT DISULFID 106 181
SQ SEQUENCE 402 AA; 45092 MW; D4FA803B4363376F CRC64;

Query Match 100.0%; Score 119; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 8.2e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFYVNLALADLCMAAFNAFNF 23
Db 69 NYFYVNLALADLCMAAFNAFNF 91

RESULT 9
NK2R_RABIT STANDARD; PRT; 384 AA.
ID NK2R_RABIT
AC P79218;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2 receptor)
DE (NK-2R) (Tachykinin receptor 2).
GN Name=TACR2; Synonyms=TACR2;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxId=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pulmonary artery;
RX MEDLINE=95182423; PubMed=7877137;
RA Aharony D., Little J., Thomas C., Powell S., Downey-Jones M.,
RA Graham A.;
RT "Isolation and characterization of neurokinin A receptor cDNAs from
RT guinea-pig lung and rabbit pulmonary artery."
RL J. Recept. Res. 14:399-421(1994).
CC -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
CC substance K (neurokinin A). It is associated with G proteins that
CC activate a phosphatidylinositol-calcium second messenger system.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL, S82448; AAB46747.1; -; mRNA.
CC InterPro, IPR000276; GPCR_Rhodopsn.
CC InterPro, IPR001681; Neurokin_receptor.

```
DR InterPro; IPR000913; NK2_receptor.
DR PANTHER; PTHR19264:SF70; NK2_receptor; 1.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01025; NEUROKININR.
DR PRINTS; PR00244; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_REC_P1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_REC_P1_2; 1.
DR G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
KW Phosphorylation; Receptor; Transducer; Transmembrane.
FT TOPO_DOM 1 32
FT TRANSMEM 33 56
FT TOPO_DOM 57 69
FT TRANSMEM 70 90
FT TOPO_DOM 91 107
FT TRANSMEM 108 129
FT TOPO_DOM 130 149
FT TRANSMEM 150 170
FT TOPO_DOM 171 196
FT TRANSMEM 197 218
FT TOPO_DOM 219 251
FT TRANSMEM 252 272
FT TOPO_DOM 273 290
FT TRANSMEM 291 310
FT TOPO_DOM 311 334
FT LIPID 324 324
FT CARBOHYD 11 11
FT CARBOHYD 19 19
FT DISULFID 106 181
SQ SEQUENCE 384 AA; 43128 MW; F35850DF224BP993 CRC64;

Query Match 96.6%; Score 115; DB 1; Length 384;
Best Local Similarity 95.7%; Pred. No. 3.3e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYFIYVNLALADLCMAAFNAFNF 23
DB 69 NYFIYVNLALADLCMAAFNAFNF 91

RESULT 10
NR2R_MESAU STANDARD; PRT; 384 AA.
AC PS1144;
DT 01-OCT-1996 (Rel. 34, Last Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2 receptor)
DE (NK-2R) (Tachykinin receptor 2).
DE Name=TACR2; Synonyms=TAC2R;
DE Mesocricetus auratus (Golden hamster).
DE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
DE Muridae; Cricetidae; Cricetinae; Mesocricetus.
DE NCBI_TaxId=10036;
DE [1]
DE NUCLEOTIDE SEQUENCE.
DE TISSUE=urinary bladder;
DE MEDLINE=94134065; PubMed=8302285;
DE Aharoni D., Little J., Thomas C., Powell S., Berry D., Graham A.;
DE "Isolation and pharmacological characterization of a hamster urinary
DE bladder neurokinin A receptor cDNA."
DE Mol. Pharmacol. 45:9-19(1994).
DE -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
DE substance K (neurokinin A). It is associated with G proteins that
DE activate a phosphatidylinositol-calcium second messenger system.
DE -1- SUBCELLULAR LOCATION: Integral membrane protein.
DE -1- MISCELLANEOUS: The rank order of affinity of this receptor to
DE tachykinins is: substance K > neuromedin K > substance P.
DE -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DE -----
DE This Swiss-Prot entry is copyright. It is produced through a collaboration
DE between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; S68899; AAC60680.1; -; mRNA.
DR HSSP; P02699; 1F88.
DR InterPro; IPR00276; GPCR_Rhodopsin.
DR InterPro; IPR001681; Neurokinin receptor.
DR InterPro; IPR000913; NK2_receptor.
DR PANTHER; PTHR19264:SF70; NK2_receptor; 1.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01025; NEUROKININR.
DR PRINTS; PR00244; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_REC_P1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_REC_P1_2; 1.
DR G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
KW Phosphorylation; Receptor; Transducer; Transmembrane.
FT TOPO_DOM 1 32
FT TRANSMEM 33 56
FT TOPO_DOM 57 69
FT TRANSMEM 70 90
FT TOPO_DOM 91 107
FT TRANSMEM 108 129
FT TOPO_DOM 130 149
FT TRANSMEM 150 170
FT TOPO_DOM 171 196
FT TRANSMEM 197 218
FT TOPO_DOM 219 251
FT TRANSMEM 252 272
FT TOPO_DOM 273 290
FT TRANSMEM 291 310
FT TOPO_DOM 311 384
FT LIPID 324 324
FT CARBOHYD 19 19
FT CARBOHYD 106 181
FT DISULFID 106 181
SQ SEQUENCE 384 AA; 43418 MW; 524BECFC3D0067F3 CRC64;

Query Match 95.8%; Score 114; DB 1; Length 384;
Best Local Similarity 91.3%; Pred. No. 4.7e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYFIYVNLALADLCMAAFNAFNF 23
DB 69 NYFIYVNLALADLCMAAFNAFNF 91

RESULT 11
NR2R_MOUSE STANDARD; PRT; 384 AA.
AC P30549;
DT 01-APR-1993 (Rel. 25, Last Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2 receptor)
DE (NK-2R) (Tachykinin receptor 2).
DE Name=Tacr2; Synonyms=Tacr2r;
DE Mus musculus (Mouse).
DE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
DE Muridae; Muridae; Murinae; Mus.
DE NCBI_TaxId=10090;
DE [1]
DE NUCLEOTIDE SEQUENCE.
DE TISSUE=Intestine;
DE MEDLINE=9213753; PubMed=1370937;
DE Sundelin J.B., Petersen P.A.;
DE Pohl J.S., Peterson P.A.;
DE "Molecular cloning of the murine substance K and substance P receptor
DE genes."
DE Eur. J. Biochem. 203:625-631(1992).
DE -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
DE substance K (neurokinin A). It is associated with G proteins that
```



```

CC activate a phosphatidylinositol-calcium second messenger system.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: The rank order of affinity of this receptor to
CC tachykinins is: substance K > neuropeptide K > substance P.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: X62933; CAA44706.1; -, mRNA.
DR PIR: S20303; S20303.
DR HSSP: P02699; 1F88.
DR Ensembl: ENSEMBL000000020081; Mus musculus.
DR MGI: MGI:98477; Tacr2.
DR GO: GO:0016021; C: integral to membrane; TAS.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR InterPro: IPR001681; Neurokin_receptor.
DR InterPro: IPR000913; NK2_receptor.
DR PANTHER: PTHR19264:SP70; NK2_receptor; 1.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PRINTS: PR00244; NEUROKININR.
DR PRINTS: PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE: PS00262; G_PROTEIN_RECP_F1_1; 1.
KW G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
KW Phosphorylation; Receptor; Transducer; Transmembrane;
KW TOPO_DOM 1 32 Extracellular (Potential).
FT TOPO_DOM 1 32 Extracellular (Potential).
FT TRANSMEM 33 56 Cytoplasmic (Potential).
FT TOPO_DOM 57 69 Cytoplasmic (Potential).
FT TRANSMEM 70 90 Extracellular (Potential).
FT TOPO_DOM 91 107 Extracellular (Potential).
FT TRANSMEM 108 129 Cytoplasmic (Potential).
FT TOPO_DOM 130 149 Cytoplasmic (Potential).
FT TRANSMEM 150 170 Extracellular (Potential).
FT TOPO_DOM 171 196 Extracellular (Potential).
FT TRANSMEM 197 218 Cytoplasmic (Potential).
FT TOPO_DOM 219 251 Cytoplasmic (Potential).
FT TRANSMEM 252 272 Extracellular (Potential).
FT TOPO_DOM 273 290 Extracellular (Potential).
FT TRANSMEM 291 310 Cytoplasmic (Potential).
FT TOPO_DOM 311 384 Cytoplasmic (Potential).
FT TOPO_DOM 324 324 S-palmitoyl cysteine (Potential).
FT LIPID 19 19 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 106 181 By similarity.
FT DISULFID 181 181 By similarity.
SQ SEQUENCE 384 AA; 43114 MW; 7937F05A2BB90759 CRC64;

Query Match 95.8%; Score 114; DB 1; Length 384;
Best Local Similarity 91.3%; Pred. No. 4.7e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Engberg S., Dmoch T.;
RA "Gerbil Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2
RA receptor) (NK-2R) (Tachykinin receptor 2).";
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -----
DR EMBL: AJ884918; CA15865.1; -, mRNA.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005886; C: plasma membrane; IEA.
DR GO: GO:0004872; F: receptor activity; IEA.
DR GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.
DR GO: GO:0004995; F: tachykinin receptor activity; IEA.
DR GO: GO:0007186; F: G-protein coupled receptor protein signalin...; IEA.
DR GO: GO:0007165; P: signal transduction; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR InterPro: IPR001681; Neurokin_receptor.
DR InterPro: IPR000913; NK2_receptor.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PRINTS: PR00244; NEUROKININR.
DR PRINTS: PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE: PS00262; G_PROTEIN_RECP_F1_1; 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 384 AA; 43305 MW; 6752F3CFAD3F67D CRC64;

Query Match 95.8%; Score 114; DB 2; Length 384;
Best Local Similarity 91.3%; Pred. No. 4.7e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Db 1 NYFIVLALADLCMAAFNAFNF 23
 69 NYFIVLALADLCMAAFNAFNF 91

RESULT 13
Q8BZV9_MOUSE PRELIMINARY; PRT; 384 AA.
 ID Q8BZV9_MOUSE PRELIMINARY; PRT; 384 AA.
 AC Q8BZV9;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Mus musculus adult male colon cDNA, K18EN full-length enriched
 DE library, clone:9030205M20 product:SUBSTANCE-K RECEPTOR, full insert
 DE sequence.
 GN Name=Tacr2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiraldi L.M., Straubli P., Suzuki R., Tomita M., Wagner L., Waeberli T.,
 RA Sakai K., Okido T., Furuno M., Momo H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita K., Gariboldi M.,
 RA Guerninch S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombarts P.,
 RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata K., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Williams L.,
 RA Wymahaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RA The PANOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama Y., Nishi K., Katsunari T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirakawa T., Hirozane T.,
 RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC EMBL; AK033443; BAC28290.1; -; mRNA.
 DR MGI; MGI:98477; Tact2.
 DR GO; GO:0016021; C:Integral to membrane; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR InterPro; IPR001681; Neurokin_receptor.
 DR InterPro; IPR000913; NK2_receptor.
 DR PANTHER; PTHR19264:SF70; NK2_receptor; 1.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PRINTS; PR01025; NEUROKININR.
 DR PRINTS; PR00244; NEUROKININR.
 DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECPT_FL_2; 1.
 KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
 KW SEQUENCE 384 AA; 43046 MW; 100B8C93547648A3 CRC64;

Query Match 95.8%; Score 114; DB 2; Length 384;
 Best Local Similarity 91.3%; Pred. No. 4.7e-09;
 Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYFINLALADLCMAAFNAAFNF 23
 |||||
 DB 69 NYFINLALADLCMAAFNAAFNF 91
 |||||
 RESULT 14
 ID NK2R RAT STANDARD; PRT; 390 AA.
 AC 116610;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Substance-K receptor (SKR) (Neurokinin A receptor) (NK-2 receptor)
 DE (NK-2R) (Tachykinin receptor 2).
 DE Name=Tact2; Synonyms=Tact2r;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=9008481; PubMed=2480781;
 RA Sasaki Y., Nakanishi S.;
 RT "Molecular characterization of rat substance K receptor and its
 RT mRNA";
 RL Biochem. Biophys. Res. Commun. 165:695-702(1989).
 CC -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
 CC substance K (neurokinin A). It is associated with G proteins that
 CC activate a phosphatidylinositol-calcium second messenger system.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: The rank order of affinity of this receptor to
 CC tachykinins is: substance K > neuromedin K > substance P.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; M31838; AAA42150.1; -; mRNA.
 DR PIR; A36737; A36737.
 DR HSP; P02699; 1P88.
 DR RGD; 3812; Tact2.
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR InterPro; IPR000913; NK2_receptor.
 DR InterPro; IPR001681; Neurokin_receptor.
 DR PANTHER; PTHR19264:SF70; NK2_receptor; 1.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PRINTS; PR01025; NEUROKININR.
 DR PRINTS; PR00244; NEUROKININR.
 DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECPT_FL_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
 KW Phosphorylation; Receptor; Transducer; Transmembrane.
 FT TOPO_DOM 1 32
 FT TRANSMEM 33 56
 FT TOPO_DOM 57 69
 FT TRANSMEM 70 90
 FT TOPO_DOM 91 107
 FT TRANSMEM 108 129
 FT TOPO_DOM 130 149
 FT TRANSMEM 150 170
 FT TOPO_DOM 171 196
 FT TRANSMEM 197 218
 FT TOPO_DOM 219 251
 Cytoplasmic (Potential).
 Extracellular (Potential).
 1 (Potential).
 Cytoplasmic (Potential).
 2 (Potential).
 Extracellular (Potential).
 3 (Potential).
 Cytoplasmic (Potential).
 4 (Potential).
 Extracellular (Potential).
 5 (Potential).
 Cytoplasmic (Potential).

```

FT TRANSMEM 252 272 6 (Potential).
FT TOPO DOM 273 290 Extracellular (Potential).
FT TRANSMEM 291 310 7 (Potential).
FT TOPO DOM 311 390 Cytoplasmic (Potential).
FT LIPID 324 324 S-palmitoyl cysteine (Potential).
FT CARBOHYD 19 19 N-linked (GlcNAc...) (Potential).
FT DISULFID 106 181 By similarity.
SQ SEQUENCE 390 AA; 43853 MW; A6C4PF0194726DB CRC64;

Query Match 95.8%; Score 114; DB 1; Length 390;
Best Local Similarity 91.3%; Pred. No. 4.7e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1 NYPIVNLATLADLCMAAFNAFNP 23
69 NYPIVNLATLADLCMAAFNAFNP 91

RESULT 15
NXLR_MOUSE STANDARD; PRT; 407 AA.
ID NXLR_MOUSE
AC P30516;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Substance-P receptor (SPR) (NK-1 receptor) (NK-1R) (Tachykinin
receptor 1).
GN Name=Tacr1; Synonyms=Tacr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC T18SUB=Intestine;
RX MEDLINE=92137253; PubMed=1370937;
RA Sundelin J.B., Provvedini D.M., Wahlestedt C.R., Laurell H.,
RA Pohl J.S., Peterson P.A.;
RT "Molecular cloning of the murine substance K and substance P receptor
genes."
RT Eur. J. Biochem. 203:625-631(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 63-290.
RC STRAIN=CBA; T18SUB=Brain, and T-cell1;
RX MEDLINE=94165478; PubMed=8120392;
RA Cook G.A., Elliott D., Metwalli A., Blum A.M., Sandor M., Lynch R.,
RA Weinstein J.V.;
RT "Molecular evidence that granuloma T lymphocytes in murine
schistosomiasis mansoni express an authentic substance P (NK-1)
receptor."
RT J. Immunol. 152:1830-1835(1994).
RL J. Immunol. 152:1830-1835(1994).
CC -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
substance P. It is probably associated with G proteins that
activate a phosphatidylinositol-calcium second messenger system.
CC The rank order of affinity of this receptor to tachykinins is:
CC substance P > substance K > neuromedin K.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X62934; CAA44707.1; -; mRNA.
DR EMBL; L27828; AAA17892.1; -; mRNA.
DR EMBL; L27826; AAA17891.1; -; mRNA.
DR PIR; S20304; S20304.
DR Ensembl; ENSMUSG00000030043; Mus musculus.
DR MGI; MGI:98475; Tacr1.
DR GO; GO:0016021; C:integral to membrane; TAS.

```

```

DR GO; GO:0048265; P:response to pain; IMP.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001681; Neurokin_receptor.
DR InterPro; IPR000046; NK1_receptor.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PRINTS; PR01024; NEUROKININR.
DR PRINTS; PR00244; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECPT_FL_2; 1.
DR G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
KM Phosphorylation; Receptor; Transducer; Transmembrane.
FT TOPO DOM 1 31 Extracellular (Potential).
FT TRANSMEM 32 54 1 (Potential).
FT TOPO DOM 55 64 Cytoplasmic (Potential).
FT TRANSMEM 65 86 2 (Potential).
FT TOPO DOM 87 106 Extracellular (Potential).
FT TRANSMEM 107 128 3 (Potential).
FT TOPO DOM 129 148 Cytoplasmic (Potential).
FT TRANSMEM 149 169 4 (Potential).
FT TOPO DOM 170 194 Extracellular (Potential).
FT TRANSMEM 195 219 5 (Potential).
FT TOPO DOM 220 248 Cytoplasmic (Potential).
FT TRANSMEM 249 270 6 (Potential).
FT TOPO DOM 271 283 Extracellular (Potential).
FT TRANSMEM 284 308 7 (Potential).
FT TOPO DOM 309 407 Cytoplasmic (Potential).
FT LIPID 322 322 S-palmitoyl cysteine (Potential).
FT CARBOHYD 14 14 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 18 18 N-linked (GlcNAc...) (Potential).
FT DISULFID 105 180 By similarity.
SQ SEQUENCE 407 AA; 46304 MW; 45186D475A659A0 CRC64;

Query Match 74.8%; Score 89; DB 1; Length 407;
Best Local Similarity 69.6%; Pred. No. 3.5e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 NYPIVNLATLADLCMAAFNAFNP 23
68 NYPIVNLATLADLCMAAFNAFNP 90

```

Search completed: January 23, 2006, 09:31:15
Job time : 61.7 secs

This Page Blank (usps)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 Seconds
(without alignments)
114.551 Million cell updates/sec

Title: US-10-501-838A-9

Perfect score: 119
Sequence: 1 NYPIVNLALADLCMAAFNAFNP 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/1aa/5.COMB.rep:*
- 2: /cgn2_6/prodata/1/1aa/6.COMB.rep:*
- 3: /cgn2_6/prodata/1/1aa/6.COMB.rep:*
- 4: /cgn2_6/prodata/1/1aa/6.COMB.rep:*
- 5: /cgn2_6/prodata/1/1aa/6.COMB.rep:*
- 6: /cgn2_6/prodata/1/1aa/6.COMB.rep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	384	1 US-08-103-170-10	Sequence 10, Appl
2	119	100.0	387	1 US-08-196-898-14	Sequence 14, Appl
3	119	100.0	387	1 US-08-760-936-14	Sequence 14, Appl
4	119	100.0	387	2 US-09-225-024-14	Sequence 14, Appl
5	119	100.0	398	2 US-09-826-509-527	Sequence 527, Appl
6	114	95.8	369	2 US-07-937-609-19	Sequence 19, Appl
7	114	95.8	369	2 US-08-029-170-19	Sequence 19, Appl
8	114	95.8	369	2 US-09-443-745-19	Sequence 19, Appl
9	89	74.8	372	1 US-07-937-609-20	Sequence 20, Appl
10	89	74.8	372	2 US-08-029-170-20	Sequence 20, Appl
11	89	74.8	372	2 US-09-443-745-20	Sequence 20, Appl
12	89	74.8	407	4 PCT-US92-06532-7	Sequence 7, Appl
13	84	70.6	407	1 US-08-380-000A-6	Sequence 6, Appl
14	79	66.4	411	1 US-07-701-935-26	Sequence 26, Appl
15	79	66.4	407	1 US-08-117-965-26	Sequence 26, Appl
16	79	66.4	407	2 US-09-826-509-525	Sequence 525, Appl
17	79	66.4	407	4 PCT-US93-06532-3	Sequence 3, Appl
18	79	66.4	465	2 US-08-090-169-1	Sequence 1, Appl
19	79	66.4	465	2 US-09-482-971-1	Sequence 1, Appl
20	79	66.4	465	2 US-09-826-509-529	Sequence 529, Appl
21	77	64.7	341	1 US-08-118-270-48	Sequence 48, Appl
22	77	64.7	341	4 PCT-US93-08528-48	Sequence 48, Appl
23	76	63.9	336	1 US-08-118-270-50	Sequence 50, Appl
24	76	63.9	336	4 PCT-US93-08528-50	Sequence 50, Appl
25	76	63.9	411	1 US-07-937-609-21	Sequence 21, Appl
26	76	63.9	411	1 US-08-029-170-21	Sequence 21, Appl
27	76	63.9	411	2 US-09-443-745-21	Sequence 21, Appl

28	76	63.9	451	2 US-08-430-286A-10	Sequence 10, Appl
29	72	60.5	440	2 US-08-430-286A-9	Sequence 9, Appl
30	72	60.5	519	2 US-09-693-746-10	Sequence 10, Appl
31	72	60.5	522	2 US-09-693-746-12	Sequence 12, Appl
32	67	56.3	612	2 US-09-270-767-32077	Sequence 32077, A
33	67	56.3	612	2 US-09-270-767-32370	Sequence 32370, A
34	67	56.3	612	2 US-09-270-767-47294	Sequence 47294, A
35	67	56.3	612	2 US-09-270-767-47587	Sequence 47587, A
36	63	52.9	1021	2 US-09-270-767-58808	Sequence 58808, A
37	63	52.9	1041	2 US-09-270-767-32619	Sequence 32619, A
38	63	52.9	1041	2 US-09-270-767-47836	Sequence 47836, A
39	63	52.9	1118	2 US-09-270-767-42703	Sequence 42703, A
40	63	52.9	1118	2 US-09-270-767-43454	Sequence 43454, A
41	63	52.9	391	2 US-09-721-870-34	Sequence 34, Appl
42	62	52.1	501	1 US-08-722-001-14	Sequence 14, Appl
43	62	52.1	501	1 US-08-467-568-9	Sequence 9, Appl
44	62	52.1	501	1 US-08-030-582-9	Sequence 9, Appl
45	62	52.1	501	1 US-09-030-582-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-103-170-10
Sequence 10, Application US/08103170
Patent No. 5883824
GENERAL INFORMATION:
APPLICANT: Yamada, Tadatsuka
APPLICANT: Gantz, Ira
TITLE OR INVENTION: Recombinant Genomic Clones Encoding
TITLE OR INVENTION: Histamine H1, H2, and H3 Receptors, Methods For Production
TITLE OR INVENTION: Thereof, and Proteins Encoded Therefrom
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/633,060
FILING DATE: 24-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-017-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 24885 OPAT UR
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE: Bos taurus
ORGANISM: Bos taurus
US-08-103-170-10
Query Match 100.0%; Score 119; DB 1; Length 384;

Beat Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFIVNLALADLCMAAFNAFNF 23
Db 69 NYFIVNLALADLCMAAFNAFNF 91

RESULT 2

US-08-196-989B-14
Sequence 14, Application US/08196989B

Patent No. 585478

GENERAL INFORMATION:

APPLICANT: MacLennan, A. John

TITLE OF INVENTION: Molecular Cloning and Expression of

TITLE OF INVENTION: G-Protein Coupled Receptors

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/196,989B

FILING DATE: 15-FEB-1994

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Lloyd, Jeff

REGISTRATION NUMBER: 35,589

REFERENCE/DOCKET NUMBER: MAC-100

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

US-08-196-989B-14

Query Match 100.0%; Score 119; DB 1; Length 387;

Beat Local Similarity 100.0%; Pred. No. 7e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFIVNLALADLCMAAFNAFNF 23
Db 69 NYFIVNLALADLCMAAFNAFNF 91

RESULT 3

US-08-760-936-14

Sequence 14, Application US/08760936

Patent No. 5856443

GENERAL INFORMATION:

APPLICANT: MacLennan, A. John

TITLE OF INVENTION: Molecular Cloning and Expression of

TITLE OF INVENTION: G-Protein Coupled Receptors

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/760,936

FILING DATE: December 6, 1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Pace, Doran R.

REGISTRATION NUMBER: 38,261

REFERENCE/DOCKET NUMBER: MAC-100C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

US-08-760-936-14

Query Match 100.0%; Score 119; DB 1; Length 387;

Beat Local Similarity 100.0%; Pred. No. 7e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFIVNLALADLCMAAFNAFNF 23
Db 69 NYFIVNLALADLCMAAFNAFNF 91

RESULT 4

US-09-225-024-14

Sequence 14, Application US/09225024

Patent No. 6518414

GENERAL INFORMATION:

APPLICANT: MacLennan, A. John

TITLE OF INVENTION: Molecular Cloning and Expression of

TITLE OF INVENTION: G-Protein Coupled Receptors

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/225,024

FILING DATE: 04-JAN-1999

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/760,936

FILING DATE: 6-DEC-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/196,989

FILING DATE: 15-FEB-1994

ATTORNEY/AGENT INFORMATION:

NAME: Pace, Doran R.

REGISTRATION NUMBER: 38,261

REFERENCE/DOCKET NUMBER: MAC-100C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULAR TYPE: protein
US-09-225-024-14

Query Match 100.0%; Score 119; DB 2; Length 387;
Best Local Similarity 100.0%; Pred. No. 76-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIYVNLALADLCMAAFNAFNF 23
Db 69 NYFIYVNLALADLCMAAFNAFNF 91

RESULT 5
US-09-826-509-527
Sequence 527, Application US/09826509
Patent No. 6806054
GENERAL INFORMATION:
APPLICANT: Lehmann-Bruhmann, Karin
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
TITLE OF INVENTION: Protein-Coupled Receptors
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: Patentin Version 2.1
SEQ ID NO 527
LENGTH: 398
TYPE: PRT
ORGANISM: Homo sapiens
US-09-826-509-527

Query Match 100.0%; Score 119; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 7,2e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIYVNLALADLCMAAFNAFNF 23
Db 69 NYFIYVNLALADLCMAAFNAFNF 91

RESULT 6
US-07-937-609-19
Sequence 19, Application US/07937609
Patent No. 5319073
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BERT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance X receptor
US-07-937-609-19

Query Match 95.8%; Score 114; DB 1; Length 369;
Best Local Similarity 91.3%; Pred. No. 3.7e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYFIYVNLALADLCMAAFNAFNF 23
Db 69 NYFIYVNLALADLCMAAFNAFNF 91

RESULT 7
US-08-029-170-19
Sequence 19, Application US/08029170
Patent No. 6169173
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-08-029-170-19

Query Match 95.8%; Score 114; DB 2; Length 369;
Best Local Similarity 91.3%; Pred. No. 3.7e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NYFIYVNLADLCMAAFNAFNF 23
Db 69 NYFIYVNLADLCMAAFNAFNF 91

RESULT 8
US-09-443-745-19
Sequence 19, Application US/09443745
Patent No. 6706493
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-09-443-745-19

Query Match 95.8%; Score 114; DB 2; Length 369;
Best Local Similarity 91.3%; Pred. No. 3.7e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NYFIYVNLADLCMAAFNAFNF 23
Db 69 NYFIYVNLADLCMAAFNAFNF 91

RESULT 9
US-07-937-609-20
Sequence 20, Application US/07937609
Patent No. 5319073
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance P receptor
US-07-937-609-20

Query Match 74.8%; Score 89; DB 1; Length 372;

Best Local Similarity 69.6%; Pred. No. 1.9e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYPIVNLADLCMAAFNAFNP 23
| | | | | : | | | | |
Db 68 NYPLVNLAFARACMAAFNTVNF 90

RESULT 10
US-08-029-170-20
Sequence 20. Application US/08029170
Patent No. 6163173
GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance P receptor
US-08-029-170-20

Query Match 74.8%; Score 89; DB 2; Length 372;
Best Local Similarity 69.6%; Pred. No. 1.9e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYPIVNLADLCMAAFNAFNP 23
| | | | | : | | | | |
Db 68 NYPLVNLAFARACMAAFNTVNF 90

RESULT 11
US-09-443-745-20

Sequence 20. Application US/09443745
Patent No. 6706493

GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance P receptor
US-09-443-745-20

Query Match 74.8%; Score 89; DB 2; Length 372;
Best Local Similarity 69.6%; Pred. No. 1.9e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYPIVNLADLCMAAFNAFNP 23
| | | | | : | | | | |
Db 68 NYPLVNLAFARACMAAFNTVNF 90

RESULT 12
PCT-US92-06532-7
Sequence 7. Application PC/TUS9206532
GENERAL INFORMATION:
APPLICANT: Krause, James E.
TITLE OF INVENTION: Human Substance P Receptor
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SD
STREET: 800 N. Lindbergh Blvd.

CITY: St. Louis
STATE: Missouri
COUNTRY: U.S.A
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06532
FILING DATE: 19920805
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: 07-24(776)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-3117
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 407 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-06532-7

Query Match 74.8%; Score 89; DB 4; Length 407;
Best Local Similarity 69.6%; Pred. No. 2.1e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYFIYVNLADLCMAAFNAAFNP 23
DB 68 NYFLVNLAFABACMAAFNTVNF 90

RESULT 13
US-08-390-000A-6
Sequence 6, Application US/08390000A
Patent No. 5985583
GENERAL INFORMATION:
APPLICANT: Sealfon, Stuart C.
TITLE OF INVENTION: Cloning and Expression of
TITLE OF INVENTION: Gonadotropin-Releasing Hormone Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,000A
FILING DATE: 17-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6923-052
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8664/9741
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 407 amino acids

TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-390-000A-6

Query Match 70.6%; Score 84; DB 1; Length 407;
Best Local Similarity 65.2%; Pred. No. 0.00012;
Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYFIYVNLADLCMAAFNAAFNP 23
DB 68 NYFLVNLAFABACMAAFNTVNF 90

RESULT 14
US-07-701-935-26
Sequence 26, Application US/07701935
Patent No. 5336595
GENERAL INFORMATION:
APPLICANT: Strader, C. D.
APPLICANT: Fong, T. M.
TITLE OF INVENTION: Method of Using Human Neurokinin-1
TITLE OF INVENTION: Receptor Short Form
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/701,935
FILING DATE: 19910517
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nicholson, William H.
REGISTRATION NUMBER: 25,147
REFERENCE/DOCKET NUMBER: 18409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-5315
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:

US-07-701-935-26

Query Match 66.4%; Score 79; DB 1; Length 311;
Best Local Similarity 65.2%; Pred. No. 0.00048;
Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 NYFIYVNLADLCMAAFNAAFNP 23
DB 68 NYFLVNLAFABACMAAFNTVNF 90

RESULT 15
US-08-117-965-26
Sequence 26, Application US/08117965
Patent No. 548486
GENERAL INFORMATION:
APPLICANT: Tung, Feng M.
APPLICANT: Cathline, Strader D.
TITLE OF INVENTION: Human Neurokinin-1 Receptor

NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,965
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 691,197
FILING DATE: 25-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Nicholson, William H.
REGISTRATION NUMBER: 25,147
REFERENCE/DOCKET NUMBER: 18387
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-5315
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 407 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
US-08-117-965-26

Query Match 66.4%; Score 79; DB 1; Length 407;
Best Local Similarity 65.2%; Pred. No. 0.00065;
Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 NYPIVNLALADLCMAAFNAFNP 23
|||:|||||:|||||
DB 68 NYPLVNLAFPAEAGMAAFNTVNP 90

Search completed: January 23, 2006, 09:35:48
Job time: 16.6 secs

This page may vary

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:06 (Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838A-9

Sequence: 1 NYPIVNLALADLCMAAFNAFNP 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	23	4	US-10-665-184-9
2	119	100.0	23	5	US-10-942-300-9
3	119	100.0	30	5	US-10-501-838A-9
4	119	100.0	30	4	US-10-665-184-31
5	119	100.0	30	5	US-10-942-300-31
6	119	100.0	30	5	US-10-501-838A-31
7	119	100.0	313	4	US-10-060-795B-7
8	119	100.0	398	4	US-09-826-509-527
9	119	100.0	398	4	US-10-225-567A-456
10	119	100.0	398	4	US-10-160-358-3
11	119	100.0	398	4	US-10-292-798-648
12	119	100.0	398	5	US-10-925-095-527
13	119	100.0	488	4	US-10-017-161-742
14	119	100.0	488	4	US-09-443-745-19
15	89	74.8	372	3	US-09-443-745-20
16	79	66.4	23	4	US-10-665-184-24
17	79	66.4	23	5	US-10-942-300-24
18	79	66.4	23	5	US-10-501-838A-24
19	79	66.4	29	4	US-10-665-184-34
20	79	66.4	29	5	US-10-942-300-34
21	79	66.4	29	5	US-10-501-838A-34
22	79	66.4	311	3	US-09-930-503-1
23	79	66.4	311	3	US-09-930-503-3
24	79	66.4	407	3	US-09-930-503-5
25	79	66.4	407	3	US-09-930-503-7
26	79	66.4	407	3	US-09-826-509-525
27	79	66.4	407	4	US-10-005-956-14

Other applications

28	79	66.4	407	4	US-10-005-956-16	Sequence 16, Appl
29	79	66.4	407	4	US-10-005-956-18	Sequence 18, Appl
30	79	66.4	407	4	US-10-005-956-20	Sequence 20, Appl
31	79	66.4	407	4	US-10-005-956-22	Sequence 22, Appl
32	79	66.4	407	4	US-10-225-567A-322	Sequence 322, App
33	79	66.4	407	5	US-10-925-095-525	Sequence 525, App
34	79	66.4	407	5	US-10-451-304-5	Sequence 5, Appl
35	79	66.4	407	5	US-10-451-304-18	Sequence 18, Appl
36	79	66.4	465	3	US-09-826-509-529	Sequence 529, App
37	79	66.4	465	3	US-10-225-567A-197	Sequence 197, App
38	79	66.4	465	5	US-10-925-095-529	Sequence 529, App
39	79	66.4	488	4	US-10-029-009-9	Sequence 9, Appl
40	79	66.4	504	5	US-10-489-425-72	Sequence 72, Appl
41	79	66.4	504	6	US-11-097-143-14043	Sequence 14043, A
42	79	66.4	505	4	US-10-029-009-21	Sequence 21, Appl
43	79	63.9	410	4	US-10-254-905-8	Sequence 8, Appl
44	76	63.9	411	3	US-09-966-782A-8	Sequence 8, Appl
45	76	63.9	411	3	US-09-443-745-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-10-665-184-9
Sequence 9, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
FILE REFERENCE: 24348-501CIP
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 9
LENGTH: 23
TYPE: PRT
ORGANISM: Homo sapiens
US-10-665-184-9

Query Match
Best Local Similarity 100.0%; Score 119; DB 4; Length 23;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYPIVNLALADLCMAAFNAFNP 23
Db 1 NYPIVNLALADLCMAAFNAFNP 23

RESULT 2
US-10-942-300-9
Sequence 9, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-503
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: US/10/942,300
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615

```

; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2.2
; SEQ ID NO 9
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
; OS-10-942-300-9

```

Query Match	100.0%	Score 119	DB 5	Length 23
Best Local Similarity	100.0%	Pred. No. 1.9e-11		
Matches 23	Conservative 0	Mismatches 0	Indels 0	Gaps 0

[illegible]

RESULT 3
US-10-501-838A-9
; Sequence 9, Application US/10501838A

```

: TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
: TITLE OF INVENTION: Biological Barrier
: FILE REFERENCE: 24348-501 NATL
: CURRENT APPLICATION NUMBER: US/10/501, 838A
: CURRENT FILING DATE: 2004-07-19
: PRIOR APPLICATION NUMBER: PCT/IB03/00968
: PRIOR FILING DATE: 2003-02-07
: PRIOR APPLICATION NUMBER: US 60/355,396
: PRIOR FILING DATE: 2002-02-07
: NUMBER OF SEQ ID NOS: 72
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 9
: LENGTH: 23
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-501-838A-9

```

Query Match	100.0%	DB 5;	Length 23;
Best Local Similarity	100.0%	Pred. No. 1.3e-11;	
Matches 23; Conservative	0;	Mismatches 0;	Gaps 0;

[illegible]

```

RESULT 4
US-10-665-184-31
Sequence 31, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
APPLICANT: Cohen, Elina
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 31
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial

```

```

? FEATURE: Penetrating peptide
? OTHER INFORMATION:
? FEATURE:
? NAME/KEY: MOD_RES
? LOCATION: (1)-(1)
? OTHER INFORMATION: ACETYLATION
? FEATURE:
? NAME/KEY: PEPTIDE
? LOCATION: (30)..(30)
? OTHER INFORMATION: wherein Xaa is Lysine
? OS-10-665-184-31

```

Query Match	100.0%	Score 119, DB 4:	length 30;
Best Local Similarity	100.0%;	Pred. No.2.5e-11;	
Matches 23;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0

Oy		1	N Y F I V N L A T A D L C M A A F N N A E N F	23
Db		1	N Y F I V N L A T A D L C M A A F N N A E N F	23

RESULT 5
US-10-942-300-31

```

: GENERAL INFORMATION:
:
: APPLICANT: Ben-Sasson, Shmuel
:
: APPLICANT: Cohen, Eliaç
:
: TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
:
: TITLE OF INVENTION: Biological Barrier

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial
;

```

```

; OTHER INFORMATION: Acylated Penetrating Peptide
;
; FEATRB:
;
; NAME/KEY: MISC_FEATURE
;

```

OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated
; OTHER INFORMATION: with a fatty acid
US-10-942-300-31

Query Match	100.0%	Score 119;	DB 5;	Length 30;
Best Local Similarity	100.0%	Pred. No.2.5e-11;		
Matches 23;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 1 NYFIVNLALADLCMAAFNNAFN 23
|||
Db 1 NYFIVNLALADLCMAAFNNAFN 23

```

RESULT 6
US-10-501-838A-31
; Sequence 31, Application US/10501838A
; Publication No. US2005025478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Saason, Shmuel A.
; APPLICANT: Cohen, Rinat
; TITLE OR INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501 NATL

```

CURRENT APPLICATION NUMBER: US/10/501,838A
CURRENT FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 31
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic; penetrating peptide
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (27)-(27)
OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
OTHER INFORMATION: groups of the lysine residue
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (30)-(30)
OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
OTHER INFORMATION: groups of the lysine residue
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (30)-(30)
OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
OTHER INFORMATION: peptide via the free amino groups of the lysine residue
US-10-501-838a-31

Query Match 100.0%; Score 119; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPIVNLALADLCMAAFNAFNF 23
DB 1 NYPIVNLALADLCMAAFNAFNF 23

RESULT 7
US-10-060-7958-7

Sequence 7, Application US/100607958
Publication No. US20030040022A1
GENERAL INFORMATION:
APPLICANT: Clevelin Olivier
APPLICANT: Bunzow, James R.
APPLICANT: Grandy, David K.
APPLICANT: Machida, Curtis A.
TITLE OF INVENTION: Dopamine Receptors and Genes
FILE REFERENCE: 90-1092-CCC
CURRENT APPLICATION NUMBER: US/10/060,7958
CURRENT FILING DATE: 2002-01-29
PRIOR APPLICATION NUMBER: 09/238977
PRIOR FILING DATE: 1999-01-27
PRIOR APPLICATION NUMBER: 08/474892
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 07/973588
PRIOR FILING DATE: 1992-11-09
PRIOR APPLICATION NUMBER: 07/438544
PRIOR FILING DATE: 1989-11-20
PRIOR APPLICATION NUMBER: 07/273372
PRIOR FILING DATE: 1988-11-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 313
TYPE: PRT
ORGANISM: Bos taurus
US-10-060-7958-7

Query Match 100.0%; Score 119; DB 4; Length 313;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPIVNLALADLCMAAFNAFNF 23
DB 69 NYPIVNLALADLCMAAFNAFNF 91

RESULT 8
US-09-826-509-527

Sequence 527, Application US/09826509
Publication No. US20030204073A1
GENERAL INFORMATION:
APPLICANT: Lehmann-Brunsmma, Karin
APPLICANT: Liaw, Chen W.
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: NO. US20030204073A1-Endogenous, Constitutively Activated Known G
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: PatentIn Version 2.1
SEQ ID NO 527
LENGTH: 398
TYPE: PRT
ORGANISM: Homo sapiens
US-09-826-509-527

Query Match 100.0%; Score 119; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPIVNLALADLCMAAFNAFNF 23
DB 69 NYPIVNLALADLCMAAFNAFNF 91

RESULT 9
US-10-225-567A-456

Sequence 456, Application US/10225567A
Publication No. US20030113798A1
GENERAL INFORMATION:
APPLICANT: Lifespan Biosciences
APPLICANT: Brown, Joseph P.
APPLICANT: Burner, Glenna C.
APPLICANT: Roush, Christine L.
TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
FILE REFERENCE: 1920-4-4
CURRENT APPLICATION NUMBER: US/10/225,567A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/257,144
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 2292
SOFTWARE: PatentIn version 3.1
SEQ ID NO 456
LENGTH: 398
TYPE: PRT
ORGANISM: Homo sapiens
US-10-225-567A-456

Query Match 100.0%; Score 119; DB 4; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPIVNLALADLCMAAFNAFNF 23
DB 69 NYPIVNLALADLCMAAFNAFNF 91

```
RESULT 10
US-10-160-358-3
; Sequence 3, Application US/10160358
; Publication No. US20030198969A1
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Bieglecki, Karyn
; APPLICANT: Cappola, Gina-Marie
; APPLICANT: Koshy, Beena
; APPLICANT: Monroe, Glen
; TITLE OF INVENTION: HAPLOTYPES OF THE TACR2 GENE
; FILE REFERENCE: TACR2.MW-0225US
; CURRENT APPLICATION NUMBER: US/10/160,358
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: PCT/US01/47394
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/247,649
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn Version 3.1
; SEQ ID NO 3
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-160-358-3

Query Match          100.0%; Score 119; DB 4; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIYNLALADLCMAAFNAFNF 23
DB 69 NYFIYNLALADLCMAAFNAFNF 91

RESULT 11
US-10-292-798-648
; Sequence 648, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 648
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-648

Query Match          100.0%; Score 119; DB 4; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIYNLALADLCMAAFNAFNF 23
DB 69 NYFIYNLALADLCMAAFNAFNF 91

RESULT 12
US-10-925-095-527
; Sequence 527, Application US/10925095
; Publication No. US20050019840A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruhsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Non-Endogenous, Constitutively Activated Known G
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/10/925,095
; CURRENT FILING DATE: 2004-08-24
; PRIOR APPLICATION NUMBER: US/09/826,509
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 527
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-925-095-527

Query Match          100.0%; Score 119; DB 5; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIYNLALADLCMAAFNAFNF 23
DB 69 NYFIYNLALADLCMAAFNAFNF 91

RESULT 13
US-10-017-161-742
; Sequence 742, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NG2G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 742
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-742

Query Match          100.0%; Score 119; DB 4; Length 488;
Best Local Similarity 100.0%; Pred. No. 5.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFIYNLALADLCMAAFNAFNF 23
DB 69 NYFIYNLALADLCMAAFNAFNF 91

RESULT 14
US-09-443-745-19
; Sequence 19, Application US/09443745
; Publication No. US20030055238A1
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; NUMBER OF SEQUENCES: 32
```


CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-09-443-745-19

Query Match 95.8%; Score 114; DB 3; Length 369;
Best Local Similarity 91.3%; Pred. No. 2.4e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYFVNLADLCMAAFNAFNF 23
DB 69 NYFVNLADLCMAAFNAFNF 91

RESULT 15
US-09-443-745-20
Sequence 20, Application US/09443745
Publication No. US20030055238A1
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance P receptor
US-09-443-745-20

Query Match 74.8%; Score 89; DB 3; Length 372;
Best Local Similarity 69.6%; Pred. No. 1.9e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYFVNLADLCMAAFNAFNF 23
DB 68 NYFVNLAFABACMAAFNTVNF 90

Search completed: January 23, 2006, 11:52:35
Job time : 61 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:56 ; Search time 5.1 Seconds
(without alignments)
45.702 Million cell updates/sec

Title: US-10-501-838A-10
Perfect score: 124
Sequence: 1 TAFDPMKMLDGVCTYVKGVQVQL 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:*
1: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pep:*
5: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pep:*
7: /cgn2_6/prodata/2/pubppa/US11_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	35.5	53	7	US-11-074-815-1 Sequence 1, Appl
2	44	35.5	1207	7	US-11-124-368A-263 Sequence 263, App
3	43	34.7	407	6	US-10-793-626-1578 Sequence 1578, App
4	42	33.9	15	7	US-11-085-812-18 Sequence 18, Appl
5	42	33.9	298	7	US-11-085-812-4 Sequence 4, Appl
6	42	33.9	330	7	US-11-085-812-2 Sequence 2, Appl
7	42	33.9	525	7	US-11-074-176-146 Sequence 146, App
8	40	32.3	58	6	US-10-667-295-218 Sequence 218, App
9	40	32.3	74	6	US-10-667-295-217 Sequence 217, App
10	40	32.3	2725	7	US-11-113-424-52 Sequence 52, Appl
11	39.5	31.9	14	6	US-10-939-890-59 Sequence 59, Appl
12	39	31.5	1001	6	US-10-467-9628-81 Sequence 81, Appl
13	38.5	31.0	192	6	US-10-821-234-1651 Sequence 1651, App
14	38.5	31.0	567	7	US-11-120-422-7 Sequence 7, Appl
15	38.5	31.0	5179	7	US-11-108-172-1068 Sequence 1068, App
16	38	30.6	262	6	US-10-467-657-8100 Sequence 8100, App
17	38	30.6	422	6	US-10-517-939-72 Sequence 72, Appl
18	38	30.6	428	6	US-10-821-234-997 Sequence 997, App
19	38	30.6	483	7	US-11-052-554A-107 Sequence 107, App
20	38	30.6	2769	7	US-11-113-424-14 Sequence 14, Appl
21	37.5	30.2	1071	6	US-10-467-657-1654 Sequence 1654, App
22	37	29.8	72	6	US-10-467-657-4798 Sequence 4798, App
23	37	29.8	122	6	US-10-793-626-820 Sequence 820, App
24	37	29.8	360	7	US-11-129-143-58 Sequence 58, Appl
25	37	29.8	879	6	US-10-770-726-78 Sequence 78, Appl

26	37	29.8	1087	7	US-11-117-169-10 Sequence 10, Appl
27	36.5	29.4	520	6	US-10-467-657-3948 Sequence 3948, App
28	36.5	29.4	1381	6	US-10-467-657-178 Sequence 178, App
29	36.5	29.4	1381	6	US-10-467-657-3726 Sequence 3726, App
30	36.5	29.4	1572	6	US-10-793-626-2906 Sequence 2906, App
31	36	29.0	121	6	US-10-467-657-3042 Sequence 3042, App
32	36	29.0	182	6	US-10-485-517-386 Sequence 386, App
33	36	29.0	319	6	US-10-467-657-1760 Sequence 1760, App
34	36	29.0	359	6	US-10-201-525-11 Sequence 11, Appl
35	36	29.0	437	6	US-10-467-657-5526 Sequence 5526, App
36	36	29.0	618	6	US-10-858-730-74 Sequence 74, Appl
37	36	29.0	757	7	US-11-052-554A-378 Sequence 378, App
38	36	29.0	803	6	US-10-821-234-1643 Sequence 42, Appl
39	36	29.0	838	7	US-11-052-554A-42 Sequence 769, App
40	36	29.0	2491	6	US-10-995-561-769 Sequence 120, App
41	35.5	28.6	352	7	US-11-156-084-120 Sequence 122, App
42	35.5	28.6	419	7	US-11-156-084-122 Sequence 349, App
43	35.5	28.6	439	7	US-11-156-084-349 Sequence 48, Appl
44	35	28.2	15	7	US-11-051-481-48 Sequence 45, Appl
45	35	28.2	66	7	US-11-051-481-45 Sequence 45, Appl

ALIGNMENTS

RESULT 1
US-11-074-815-1
; Sequence 1, Application US/11074815
; Publication No. US20050275799A1
; GENERAL INFORMATION:
; APPLICANT: Ocular Sciences, Inc.
; TITLE OF INVENTION: Contact lenses, package systems, and method for promoting a
; TITLE OF INVENTION: healthy epithelium of an eye
; FILE REFERENCE: D-4118
; CURRENT APPLICATION NUMBER: US/11/074,815
; PRIOR FILING DATE: 2005-03-08
; PRIOR APPLICATION NUMBER: 60/551,834
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-074-815-1

Query Match 35.5%; Score 44; DB 7; Length 53;
Best Local Similarity 46.2%; Pred. No. 0.9; 3; Indels 0; Gaps 0;
Matches 6; Conservative 4; Mismatches

Qy 10 DGVCTYVKGVQV 22
Db 17 DGVCTYVKGVQV 29

RESULT 2
US-11-124-368A-263
; Sequence 263, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CU001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09

NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-124-368A-263

Query Match 35.5%; Score 44; DB 7; Length 1207;
Best Local Similarity 46.2%; Pred. No. 27;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 GVCCTYKGVQOQ 22
DB 987 GVCCTYKGVQOQ 999

RESULT 3
US-10-793-626-1578
; Sequence 1578, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; NUMBER OF SEQ ID NOS: 1999-11-09
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1578
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1578

Query Match 34.7%; Score 43; DB 6; Length 407;
Best Local Similarity 45.5%; Pred. No. 12;
Matches 10; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 ADEPKMLDGVCTYKGVQOQ 23
DB 107 ADEPKMLDGVCTYKGVQOQ 128

RESULT 4
US-11-085-812-18
; Sequence 18, Application US/11085812
; Publication No. US20050261186A1
; GENERAL INFORMATION:
; APPLICANT: Marchionni, Mark
; TITLE OF INVENTION: NRG-2 NUCLEIC ACID MOLECULES,
; FILE REFERENCE: POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS
; CURRENT APPLICATION NUMBER: US/11/085,812
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/206,495
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: derived from Rattus norvegicus and Homo sapiens
US-11-085-812-18

Query Match 33.9%; Score 42; DB 7; Length 15;

Best Local Similarity 54.5%; Pred. No. 0.47;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 GVCCTYKGVQOQ 21
DB 3 GVCCTYKGVQOQ 13

RESULT 5
US-11-085-812-4
; Sequence 4, Application US/11085812
; Publication No. US20050261186A1
; GENERAL INFORMATION:
; APPLICANT: Marchionni, Mark
; TITLE OF INVENTION: NRG-2 NUCLEIC ACID MOLECULES,
; FILE REFERENCE: POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS
; CURRENT APPLICATION NUMBER: US/11/085,812
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/206,495
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-085-812-4

Query Match 33.9%; Score 42; DB 7; Length 298;
Best Local Similarity 54.5%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 GVCCTYKGVQOQ 21
DB 265 GVCCTYKGVQOQ 275

RESULT 6
US-11-085-812-2
; Sequence 2, Application US/11085812
; Publication No. US20050261186A1
; GENERAL INFORMATION:
; APPLICANT: Marchionni, Mark
; TITLE OF INVENTION: NRG-2 NUCLEIC ACID MOLECULES,
; FILE REFERENCE: POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS
; CURRENT APPLICATION NUMBER: US/11/085,812
; PRIOR FILING DATE: 2005-03-21
; PRIOR APPLICATION NUMBER: US 60/206,495
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-085-812-2

Query Match 33.9%; Score 42; DB 7; Length 330;
Best Local Similarity 54.5%; Pred. No. 14;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 GVCCTYKGVQOQ 21
DB 265 GVCCTYKGVQOQ 275

RESULT 7
US-11-074-176-146
; Sequence 146, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:

```

/ APPLICANT: Klaenhammer, Todd R.
/ APPLICANT: Russell, William M.
/ APPLICANT: Altermann, Eric
/ APPLICANT: McAlliffe, Olivia
/ APPLICANT: Perill, Andrea Accarate
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding
/ TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
/ FILE REFERENCE: 5051-694
/ CURRENT APPLICATION NUMBER: US/11/074,176
/ PRIOR FILING DATE: 2005-03-07
/ PRIOR APPLICATION NUMBER: 60/551,161
/ PRIOR FILING DATE: 2004-03-08
/ NUMBER OF SEQ ID NOS: 381
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 146
/ LENGTH: 525
/ TYPE: PRT
/ ORGANISM: Lactobacillus acidophilus
/ US-11-074-176-146

```

```

Query Match      33.9%; Score 42; DB 7; Length 525;
Best Local Similarity 36.4%; Pred. No. 23;
Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      2 AFDPMALDGVCTYVKGVOYL 23
DB      262 ATGINKMLDMQSYIEQKQFV 283

```

```

RESULT 8
US-10-667-295-218
/ Sequence 218, Application US/10667295
/ Publication No. US20050257293A1
/ GENERAL INFORMATION:
/ APPLICANT: Maciel, Peter
/ TITLE OF INVENTION: BIOLOGICAL CONTAINMENT SYSTEM
/ FILE REFERENCE: 11696-047001
/ CURRENT APPLICATION NUMBER: US/10/667,295
/ CURRENT FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: US 60/411,823
/ PRIOR FILING DATE: 2002-09-17
/ NUMBER OF SEQ ID NOS: 263
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 218
/ LENGTH: 58
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(58)
/ OTHER INFORMATION: Ceres Seq. ID no. 13571377
/ US-10-667-295-218

```

```

Query Match      32.3%; Score 40; DB 6; Length 58;
Best Local Similarity 40.0%; Pred. No. 4.3;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      7 KMLDGVCTYVKGVOQ 21
DB      28 KLMKTCSTYIKSLQR 42

```

```

RESULT 9
US-10-667-295-217
/ Sequence 217, Application US/10667295
/ Publication No. US20050257293A1
/ GENERAL INFORMATION:
/ APPLICANT: Maciel, Peter
/ TITLE OF INVENTION: BIOLOGICAL CONTAINMENT SYSTEM
/ FILE REFERENCE: 11696-047001
/ CURRENT APPLICATION NUMBER: US/10/667,295
/ CURRENT FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: US 60/411,823

```

```

/ PRIOR FILING DATE: 2002-09-17
/ NUMBER OF SEQ ID NOS: 263
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 217
/ LENGTH: 74
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: (1)...(74)
/ OTHER INFORMATION: Ceres Seq. ID no. 13571376
/ US-10-667-295-217

```

```

Query Match      32.3%; Score 40; DB 6; Length 74;
Best Local Similarity 40.0%; Pred. No. 5.6;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      7 KMLDGVCTYVKGVOQ 21
DB      44 KLMKTCSTYIKSLQR 58

```

```

RESULT 10
US-11-113-424-52
/ Sequence 52, Application US/11113424
/ Publication No. US20050260713A1
/ GENERAL INFORMATION:
/ APPLICANT: Gangoli et al.
/ TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-225
/ CURRENT APPLICATION NUMBER: US/11/113,424
/ CURRENT FILING DATE: 2005-04-21
/ PRIOR APPLICATION NUMBER: 60/256,704
/ PRIOR FILING DATE: 2000-12-19
/ PRIOR APPLICATION NUMBER: 60/311,590
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: 60/257,314
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 60/311,613
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: 60/315,617
/ PRIOR FILING DATE: 2001-08-29
/ PRIOR APPLICATION NUMBER: 60/307,506
/ PRIOR FILING DATE: 2001-07-24
/ PRIOR APPLICATION NUMBER: 60/322,358
/ PRIOR FILING DATE: 2001-09-14
/ PRIOR APPLICATION NUMBER: 60/294,075
/ PRIOR FILING DATE: 2001-05-29
/ PRIOR APPLICATION NUMBER: 60/288,153
/ PRIOR FILING DATE: 2001-05-02
/ NUMBER OF SEQ ID NOS: 190
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 52
/ LENGTH: 2725
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-11-113-424-52

```

```

Query Match      32.3%; Score 40; DB 7; Length 2725;
Best Local Similarity 63.6%; Pred. No. 2.9e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 TAFDPMALDVG 11
DB      455 TQDFVTKLMDG 465

```

```

RESULT 11
US-10-939-890-59
/ Sequence 59, Application US/10939890
/ Publication No. US20050250700A1
/ GENERAL INFORMATION:
/ APPLICANT: Sato, Aaron K.

```

```

; APPLICANT: Sexton, Daniel J.
; APPLICANT: Dransfield, Daniel T.
; APPLICANT: Lachner, Robert C.
; APPLICANT: Arbogast, Christophe
; APPLICANT: Buesat, Philippe
; APPLICANT: Fan, Hong
; APPLICANT: Khurana, Sudha
; APPLICANT: Linder, Karen E.
; APPLICANT: Marinelli, Edmund R.
; APPLICANT: Nanjappa, Palanappa
; APPLICANT: Nunn, Adrian D.
; APPLICANT: Pillai, Radhakrishna
; APPLICANT: Pochon, Sibylle
; APPLICANT: Ramalingam, Kondareddiar
; APPLICANT: Shrivastava, Ajay
; APPLICANT: Song, Bo
; APPLICANT: Swenson, Rolf E.
; APPLICANT: Von Wronski, Mathew A.
; TITLE OF INVENTION: KDR AND VEGF/KDR BINDING PEPTIDES
; FILE REFERENCE: D0617, 70014US90
; CURRENT APPLICATION NUMBER: US/10/939,890
; CURRENT FILING DATE: 2004-09-13
; PRIOR APPLICATION NUMBER: US 10/661,156
; PRIOR FILING DATE: 2003-09-11
; PRIOR APPLICATION NUMBER: US 10/382,082
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: PCT/US03/06731
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: US 60/440,411
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/360,851
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 883
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Library Isolate
;
; US-10-939-890-59
;
Query Match          31.9%; Score 39.5; DB 6; Length 14;
Best Local Similarity 72.7%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 1; Gaps 1;

Qy          4 DEN-KMLDGVVC 13
Db          1 DENCKMIDGFC 11

RESULT 12
; US-10-467-962B-81
; Sequence 81, Application US/10467962B
; Publication No. US20050246784A1
; GENERAL INFORMATION:
; APPLICANT: Plesch, Gunmar
; APPLICANT: Blaue, Astrid
; APPLICANT: Daeschner, Klaus
; APPLICANT: Klein, Mathieu
; TITLE OF INVENTION: Identification of Herbicidally Active Substances
; FILE REFERENCE: 2000 857
; CURRENT APPLICATION NUMBER: US/10/467,962B
; CURRENT FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: PCT/EP02/01466
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 81
; LENGTH: 1001
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-10-467-962B-81
```

```

Query Match          31.5%; Score 39; DB 6; Length 1001;
Best Local Similarity 37.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy          7 KMLDGVCTYKGVQY 22
Db          223 QVLEGLTRLNGVRQF 238

RESULT 13
; US-10-821-234-1651
; Sequence 1651, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crafin, Birgit
; APPLICANT: Andarmat, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: Pf SEQ_genes Version 1.0
; SEQ ID NO 1651
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-821-234-1651
;
Query Match          31.0%; Score 38.5; DB 6; Length 192;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

Qy          4 DENKMLDGVCTYKGVQY 21
Db          171 DIRKFLDGIYSEKGTVOQ 189

RESULT 14
; US-11-120-422-7
; Sequence 7, Application US/11120422
; Publication No. US20050287646A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kehama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS BOVISIMILIS HYALURONAN SYNTHASE GENE AND EXPRESSION
; TITLE OF INVENTION: IN BACILLUS SUBTILIS
; FILE REFERENCE: 3554, 049
; CURRENT APPLICATION NUMBER: US/11/120,422
; CURRENT FILING DATE: 2005-05-02
; PRIOR APPLICATION NUMBER: US/09/879,959
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Chlorella virus PBCV-1
; US-11-120-422-7
;
Query Match          31.0%; Score 38.5; DB 7; Length 567;
Best Local Similarity 36.8%; Pred. No. 90;
Matches 7; Conservative 5; Mismatches 2; Indels 5; Gaps 1;
```


THE UNIVERSITY OF CHICAGO

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:04:28 ; Search time 69.2 Seconds
(without alignments)
146.036 Million cell updates/sec

Title: US-10-501-838a-10

Sequence: 1 TAFDFNKMIDGVCTYKGVQVYL 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: geneseq_21.*
2: geneseqp1908.*
3: geneseqp2000.*
4: geneseqp2001.*
5: geneseqp2002.*
6: geneseqp2003.*
7: geneseqp2003b.*
8: geneseqp2004.*
9: geneseqp2005.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	124	100.0	23	ADBI6895 Chlamydia
2	124	100.0	23	ABO88227 Chlamydia
3	124	100.0	91	AAY35335 Amino aci
4	101	81.5	90	AAY37513 Protein w
5	94	75.8	86	ADR99475 Chlamydia
6	50	40.3	520	ABM91346 M. xanthu
7	48.5	39.1	384	ADN46751 Thermococ
8	48.5	39.1	699	ADN46751 Bacterial
9	48	38.7	44	AAY59060 Tek recep
10	48	38.7	53	ABO88420 Mature mo
11	48	38.7	53	ABO88420 Mature mo
12	48	38.7	53	ABO88422 Mature mo
13	48	38.7	53	ABO88421 Mature mo
14	48	38.7	53	ABO88419 Mature mo
15	48	38.7	52	ABO88449 Dog epide
16	47	37.9	285	ABU24753 Protein e
17	47	37.9	507	ADN55618 Rat Prote
18	47	37.9	507	ADN55622 Rat Prote
19	47	37.9	509	ADQ97592 Mouse can
20	47	37.9	516	ABO63689 Drosophi
21	46.5	37.5	716	ABM69326 Phototrab
22	46	37.1	52	AAR24937 Canine ce
23	46	37.1	52	ADJ32585 Dog epide
24	46	37.1	9	ABO86450 Cat epide

25	46	37.1	53	2	AAM34464	Aaw34464 Chimeric
26	46	37.1	53	2	AAM34466	Aaw34466 Chimeric
27	46	37.1	53	2	AAM50137	Aaw50137 Epidermal
28	46	37.1	53	2	AAM50139	Aaw50139 Epidermal
29	46	37.1	53	2	AAY01791	Aay01791 Epithelia
30	46	37.1	53	2	AAY01793	Aay01793 Epithelia
31	46	37.1	68	4	AAG74675	Aag74675 Human col
32	46	37.1	112	6	ABP72538	Abp72538 Multivale
33	46	37.1	9	6	ABE39483	Abp39483 L. pneumo
34	46	37.1	307	9	ABE36055	Abp36055 L. pneumo
35	46	37.1	319	9	ADG42157	Adg42157 Bacterial
36	45	36.3	52	8	ADJ32584	Adj32584 Cat epide
37	45	36.3	53	1	AAP81343	Aap81343 Sequence
38	45	36.3	53	1	AAP83165	Aap83165 Sequence
39	45	36.3	53	2	AAR08004	Aar08004 Modified
40	45	36.3	53	2	AAR32857	Aar32857 Epidermal
41	45	36.3	53	2	AAR32858	Aar32858 BGP-A16
42	45	36.3	53	2	AAR32859	Aar32859 BGP-B12
43	45	36.3	53	2	AAR32860	Aar32860 BGP-W13
44	45	36.3	53	2	AAR32865	Aar32865 BGP-B11
45	45	36.3	53	2	AAR32866	Aar32866 BGP-B11

ALIGNMENTS

RESULT 1

ADBI6895
ID ADBI6895 standard; peptide; 23 AA.

ADBI6895;

20-NOV-2003 (first entry)

Chlamydia pneumoniae CPN0710/C penetrating peptide 10.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;

infertility; hormone; vitamin deficiency; neurodegenerative;

cardiovascular; haematological; endocrine disorder; obesity;

neoplastic disease; neuroprotective; cardiatic; antiarteriosclerotic;

osteopathic; cytosaratic; nootropic.

Chlamydia pneumoniae.

MO2003066853-A2.

07-FEB-2003; 2003MO-IB000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen B;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating

or preventing e.g. endocrine disorders.

Claim 2; Page 14; 60pp; English.

This invention relates to a novel peptide sequences capable of

translocating across a biological barrier. Furthermore, it refers to

methods that use these peptides to facilitate penetration of a

biologically active effector molecule such as a drug or other therapeutic

agent across biological barriers e.g. epithelial or endothelial cells

sealed by tight junctions. This peptide is derived from a bacterial

toxin, an integral membrane or extracellular protein and can comprise an

anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin

or enzyme. The effector molecule, however, can comprise for example

insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from CPN0710/C of
 CC Chlamydia pneumoniae and is penetrating peptide 10 of the invention.

XX Sequence 23 AA;

Query Match 100.0%; Score 124; DB 6; Length 23;

Best Local Similarity 100.0%; Pred. No. 4,4e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCITYKGVQGYL 23
 |||||
 Db 1 TAFDFNKMLDGVCITYKGVQGYL 23

RESULT 2
 AEB08227

ID AEB08227 standard; peptide; 23 AA.

XX AC AEB08227;

XX DT 25-AUG-2005 (first entry)

XX Chlamydia pneumoniae CPN0710/C penetrating peptide 10, SEQ ID NO: 10.

XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; virocidic;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antirheumatism; antirheumatic; cyostatic;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.

XX Chlamydia pneumoniae.

XX US2005136103-A1.

XX 23-JUN-2005.

XX 16-SEP-2004; 2004US-00942300.

XX 17-SEP-2003; 2003US-00664989.

XX 17-SEP-2003; 2003US-00665184.

XX 17-SEP-2003; 2003US-0503615P.

XX PA (BENS/) BEN-SASSON S A.

XX PA (COHEN/) COHEN E.

XX Ben-Sasson SA, Cohen E;

XX WPI; 2005-444089/45.

PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Clam 46; SEQ ID NO 10; 59pp; English.

CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transmembrane delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Chlamydia pneumoniae
 CC CPN0710/C penetrating peptide. This sequence is used in the effective
 CC translocation of aminoglycoside antibiotics and antifungal agents across
 CC an epithelial barrier.

XX Sequence 23 AA;

Query Match 100.0%; Score 124; DB 9; Length 23;

Best Local Similarity 100.0%; Pred. No. 4,4e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCITYKGVQGYL 23
 |||||
 Db 1 TAFDFNKMLDGVCITYKGVQGYL 23

RESULT 3
 AAY35335

ID AAY35335 standard; protein; 91 AA.

XX AC AAY35335;

XX DT 17-OCT-2003 (revised)

XX DT 13-SEP-1999 (first entry)

XX Amino acid sequence of a Chlamydia pneumoniae protein.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
 XX neutralising epitope.

XX Chlamydia pneumoniae.

XX WO9927105-A2.

XX 03-JUN-1999.

XX 20-NOV-1998; 98WO-1B001890.

XX 21-NOV-1997; 97FR-00014673.

XX PR 04-NOV-1998; 98US-0107078P.

XX (GENST) GENSET.

XX Griffiths R;

XX WPI; 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae.

XX Page 1141; Disclosure; 1912pp; English.

PS AAY34584-Y35879 represent the proteins encoded by all the open reading
 XX frames in the complete genome (see AAY31990) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis,
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,
 CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -OCT-2003 to standardise OS field)

XX Sequence 91 AA:

SO Query Match 100.0%; Score 124; DB 2; Length 91;
 Best Local Similarity 100.0%; Pred. No. 2, 1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAFDFNKMIDGCTCYKGVQVYL 23
 |||||
 DB 15 TAFDFNKMIDGCTCYKGVQVYL 37

RESULT 4
 AAY37513 standard; protein; 90 AA.

XX AAY37513;
 XX 07-OCT-1999 (first entry)
 XX Protein which is specific to Chlamydia trachomatis.
 XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 XX paratrachoma; inclusion conjunctivitis; genital disease; peritrichitis;
 XX nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 XX bartolinitis; pneumopathy; venereal lymphogranulomatosis.
 XX Chlamydia trachomatis.
 XX OS
 XX MO928475-A2.
 XX PN
 XX 10-JUN-1999.
 XX 27-NOV-1998; 98WC-IB001939.
 XX 28-NOV-1997; 97FR-00015041.
 XX 17-DEC-1997; 97FR-00016034.
 XX 04-NOV-1998; 98US-0107077P.
 XX (GENSET) GENSET.
 XX PI
 XX Giffels R;
 XX WPI; 1999-371125/31.
 XX Genome sequence of Chlamydia trachomatis.
 XX PT
 XX Disclosure; Page 1186-1187; 1755pp; English.

CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
 CC of Chlamydia trachomatis (see AAY01425). The polypeptides can be used as
 CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 CC can also be used to control growth of the microorganism. Chlamydia
 CC trachomatis is responsible for a large number of diseases, e.g. eye
 CC diseases such as conventional trachoma, nonendemic trachoma,
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis, and
 CC peritrichitis, bartolinitis; pneumopathy in breast feeding infants; and
 CC venereal lymphogranulomatosis. The polypeptides of the invention may be

CC of use in treating these diseases

XX Sequence 90 AA:

SO Query Match 81.5%; Score 101; DB 2; Length 90;
 Best Local Similarity 78.3%; Pred. No. 8, 1e-08;
 Matches 18; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 TAFDFNKMIDGCTCYKGVQVYL 23
 :|||:|
 DB 14 SAFNFNKMIDGCTCYKGVQVYL 36

RESULT 5
 ADR9475 standard; protein; 86 AA.

XX ADR9475;
 XX 18-NOV-2004 (first entry)
 XX Chlamydia psittaci protein #7 for treating Chlamydia infections.
 XX antibacterial; antiinflammatory; antiarteriosclerotic; gene therapy;
 XX diagnosis; sexually transmitted disease; respiratory disease; pneumonia;
 XX bronchitis.
 XX Chlamydia psittaci.
 XX OS
 XX MO2004074318-A2.
 XX PN
 XX 02-SEP-2004.
 XX 24-FEB-2004; 2004WC-IB000902.
 XX 24-FEB-2003; 2003US-0448879P.
 XX (INSP) INST PASTEUR.
 XX (CNRS) CENT NAT RECH SCT.
 XX PI
 XX Dautry-Varvat A, Subtil-Sande A;
 XX WPI; 2004-652913/63.
 XX N-PSDB; ADR9474.
 XX New purified secreted Chlamydia polypeptides and encoding polynucleotides
 XX for diagnosing, preventing or treating Chlamydia infections, such as
 XX sexually transmitted diseases, respiratory diseases or arteriosclerosis.
 XX Disclosure; SEQ ID NO 120; 210pp; English.

PS The invention relates to a purified secreted Chlamydia polypeptide or its
 CC fragment. The Chlamydia polypeptide is homologous to one or more
 CC Chlamydia pneumoniae proteins selected from CPN0104, CPN0206, CPN0210,
 CC CPN0399, CPN0405, CPN0443, CPN0480, CPN0489, CPN0490, CPN0497, CPN0522,
 CC CPN0556, CPN0588, CPN0589, CPN0595, CPN0671, CPN0672, CPN0681, CPN0712,
 CC CPN0725, CPN0729, CPN0746, CPN0755, CPN0764, CPN0770,
 CC CPN0774, CPN0792, CPN0853, CPN0859, CPN0879, CPN0906, CPN0939, CPN1002,
 CC CPN1005, CPN1007, CPN1019, CPN1020, CPN1032 and CPN1058. Alternatively,
 CC the purified secreted Chlamydia polypeptide is homologous to one or more
 CC Chlamydia trachomatis proteins selected from CT87, CT476, CT550,
 CC CT606.1, CT610, CT642, CT652.1, CT664, CT718, CT763, CT845 and CT848. The
 CC Chlamydia polypeptide is also homologous to one or more Chlamydia
 CC psittaci proteins selected from Ps10330, Ps10379, Ps10595, Ps10648,
 CC Ps10671, Ps10705, Ps10710, Ps10761, Ps10774, Ps11002, Ps11005, Ps11022
 CC and Ps11058. The composition and methods are useful for diagnosing,
 CC preventing and treating Chlamydia infections, such as sexually
 CC transmitted diseases, respiratory diseases (e.g. pneumonia or bronchitis)
 CC or arteriosclerosis. The may also be used in drug screening purposes. This
 CC sequence corresponds to one of the Chlamydia proteins of the invention.

XX Sequence 86 AA;

KM pathogen tolerance; pest tolerance; plant disease resistance;
 KM cell cycle pathway modification; plant growth regulator;
 KM homologous recombination; seed oil yield; protein yield; carbohydrate;
 KM nitrogen; phosphorus; photosynthesis; lignin; galactomanan;
 KM bacterial polypeptide.
 OS Bacteria.
 XX US200323675-A1.
 XX 18-DEC-2003.
 XX 20-FEB-2003; 2003US-00369493.
 XX 21-FEB-2002; 2002US-0360039P.
 XX (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.
 XX (CHEN/) CHEN X.
 XX (GOLD/) GOLDMAN B S.
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX Claim 1; SEQ ID NO 21251; 122pp; English.
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition. Improved lignin production or improved galactomanan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX Sequence 699 AA;
 XX
 XX Query Match 39.1%; Score 48.5; DB 8; Length 699;
 XX Best Local Similarity 47.8%; Pred. No. 1.3e+02;
 XX Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;
 XX
 XX 1 TAPDFNMLDGVCTYVKGVOYL 23
 XX 268 TLFQNLNGQLD--TFVNVNVEOYL 287
 XX
 XX RESULT 9
 XX AAY59060
 XX ID AAY59060 standard; protein; 44 AA.
 XX XX
 XX AC AAY59060;

XX 07-MAR-2000 (first entry)
 XX Tek receptor EGF sequence.
 XX Receptor tyrosine kinase protein; angiogenesis; cardiogenesis; mouse;
 KM tumorigenesis; EGF.
 OS Mus musculus.
 XX US5998187-A.
 XX 07-DEC-1999.
 XX 23-APR-1997; 97US-00838957.
 XX 30-JUL-1992; 92US-00921795.
 XX 29-APR-1994; 94US-00235408.
 XX 20-JUL-1994; 94US-00278089.
 XX (MOUN) MOUNT SINAI HOSPITAL CORP.
 XX Yamaguchi TP, Breitman J, Dumont DJ, Rossant J, Breitman ML;
 XX WPI; 2000-052545/04.
 XX Novel receptor tyrosine kinase protein involved in angiogenesis,
 PT cardiogenesis and tumorigenesis.
 XX Example 7; Col 93-94; 122pp; English.
 XX The invention provides novel receptor tyrosine kinase protein sequences
 CC (AAY59046-48) and nucleic acids encoding the polypeptides. The isolated
 CC and purified polypeptides are useful for studying the developmental
 CC expression of a receptor tyrosine kinase protein which may have a role in
 CC angiogenesis, cardiogenesis and tumorigenesis
 XX Sequence 44 AA;
 XX
 XX Query Match 38.7%; Score 48; DB 3; Length 44;
 XX Best Local Similarity 53.8%; Pred. No. 7.4;
 XX Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 XX
 XX 10 DGVCTYVKGVOY 22
 XX 17 DGVCTYVKGVOY 29
 XX
 XX RESULT 10
 XX AEB86420
 XX ID AEB86420 standard; protein; 53 AA.
 XX AEB86420;
 XX AC AEB86420;
 XX 06-OCT-2005 (first entry)
 XX Mature modified epidermal growth factor SEQ ID NO:13.
 XX protein engineering; epidermal growth factor; cytosolic; antimicrobial;
 KM vasotropic; antineumatic; antiarthritic; antiinflammatory;
 KM proliferative disease; cancer; immune disorder; infectious disease;
 KM vascular disease.
 OS Homo sapiens.
 OS Synthetic.
 XX EP1557429-A1.
 XX 27-JUL-2005.
 XX 23-JAN-2004; 2004EP-00001454.
 XX 23-JAN-2004; 2004EP-00001454.

XX AEB86422;
 AC
 XX
 DT 06-OCT-2005 (first entry)
 XX
 DE Mature modified epidermal growth factor SEQ ID NO:15.
 XX
 KM protein engineering; epidermal growth factor; cytostatic; antimicrobial;
 KM vasotrophic; antineumatic; antiairchitic; antinflammatory;
 KM proliferative disease; cancer; immune disorder; infectious disease;
 KM vascular disease.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN EPI557429-A1.
 XX
 PD 27-JUL-2005.
 XX
 PF 23-JAN-2004; 2004EP-00001454.
 XX
 PR 23-JAN-2004; 2004EP-00001454.
 XX
 PA (VECT-) VECTRON THERAPEUTICS AG.
 XX
 PI Konferman R.
 XX
 DR WPI; 2005-573617/59.
 XX
 DR N-PSDB; AEB86437.
 XX
 PT New polynucleotide encoding a mature modified epidermal growth factor,
 PT for producing a medicament for therapy and for diagnosing e.g.
 PT proliferative diseases, immune diseases, and vascular diseases.
 XX
 PS Claim 1; SEQ ID NO 15; 50bp; English.

CC The invention relates to a polynucleotide (I) chosen from a
 CC polynucleotides encoding a mature modified epidermal growth factor (EGF)
 CC having a 53 amino acid sequence of AEB86408 to AEB86422, where the
 CC polynucleotide has a 159 base pair sequence of AEB86423 to AEB86437, and
 CC encodes mature modified EGF. Also described: (1) a vector (VI) containing
 CC (1); (2) a host cell (H1) genetically engineered with (1) or (VI); (3)
 CC producing a modified EGF encoded by (1); (4) producing (M1) cells
 CC recovering the modified EGF encoded by (1); (4) producing (M1) cells
 CC capable of expressing modified EGF, comprising genetically engineering
 CC cells in vitro with (VI), where the modified EGF is encoded by (1); (5) a
 CC modified EGF (II) having the amino acid sequence encoded by (1) or (II) or
 CC obtainable by (M1); (6) a composition (C1) comprising one or more (II) or
 CC a fusion polypeptide and one or more further component chosen from
 CC liposomes, virosomes, microspheres, niosomes, dendrimers, stabilizers,
 CC buffers, excipients and additives; (7) producing (M2) a modified binding
 CC polypeptide, which is suitable for site-directed coupling, comprising
 CC modifying a polynucleotide encoding the binding polypeptide; and (8) use
 CC of a modified binding polypeptide or fusion polypeptide (III) producible
 CC by (M3) for the manufacture of a medicament or diagnostic for the
 CC prevention, treatment or diagnosis of a disease, which is characterized
 CC by an increased or decreased amount of at least one binding partner of
 CC the binding polypeptide in diseased tissue or cells involved in the
 CC disease. (I), (II) or its fusion polypeptide, or (C1) are useful for
 CC producing a medicament for the therapy of proliferative diseases, immune
 CC diseases, infectious diseases, vascular diseases, rheumatoid diseases,
 CC and diseases in which cells in or adjacent to the disease site show an
 CC increased expression of EGF, or for the diagnosis of the diseases. (III)
 CC is useful for manufacturing a medicament or diagnostic for the
 CC prevention, treatment or diagnosis of the diseases. (II) is useful for
 CC site-specific coupling e.g., targeting ligands and modified human EGF and
 CC its fragments, suitable for site-specific coupling. The present sequence
 CC represents a modified mature EGF amino acid sequence, which was modified
 CC from human EGF.
 XX
 XX Sequence 53 AA;
 SQ

Query Match 38.7%; Score 48; DB 9; Length 53;

Best Local Similarity 53.8%; Pred. No. 9.1;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 10 DGVCYVKGVOOY 22
 |||||:::
 17 DGVCYVKGVOOY 29

DB
 RESULT 13
 AEB86421
 ID AEB86421 standard; protein; 53 AA.
 XX
 AC AEB86421;
 XX
 DT 06-OCT-2005 (first entry)
 XX
 DE Mature modified epidermal growth factor SEQ ID NO:14.
 XX
 KM protein engineering; epidermal growth factor; cytostatic; antimicrobial;
 KM vasotrophic; antineumatic; antiairchitic; antinflammatory;
 KM proliferative disease; cancer; immune disorder; infectious disease;
 KM vascular disease.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN EPI557429-A1.
 XX
 PD 27-JUL-2005.
 XX
 PF 23-JAN-2004; 2004EP-00001454.
 XX
 PR 23-JAN-2004; 2004EP-00001454.
 XX
 PA (VECT-) VECTRON THERAPEUTICS AG.
 XX
 PI Konferman R.
 XX
 DR WPI; 2005-573617/59.
 XX
 DR N-PSDB; AEB86436.
 XX
 PT New polynucleotide encoding a mature modified epidermal growth factor,
 PT for producing a medicament for therapy and for diagnosing e.g.
 PT proliferative diseases, immune diseases, and vascular diseases.
 XX
 PS Claim 1; SEQ ID NO 14; 50bp; English.

CC The invention relates to a polynucleotide (I) chosen from a
 CC polynucleotides encoding a mature modified epidermal growth factor (EGF)
 CC having a 53 amino acid sequence of AEB86408 to AEB86422, where the
 CC polynucleotide has a 159 base pair sequence of AEB86423 to AEB86437, and
 CC encodes mature modified EGF. Also described: (1) a vector (VI) containing
 CC (1); (2) a host cell (H1) genetically engineered with (1) or (VI); (3)
 CC producing a modified EGF encoded by (1); (4) producing (M1) cells
 CC recovering the modified EGF encoded by (1); (4) producing (M1) cells
 CC capable of expressing modified EGF, comprising genetically engineering
 CC cells in vitro with (VI), where the modified EGF is encoded by (1); (5) a
 CC modified EGF (II) having the amino acid sequence encoded by (1) or
 CC obtainable by (M1); (6) a composition (C1) comprising one or more (II) or
 CC a fusion polypeptide and one or more further component chosen from
 CC liposomes, virosomes, microspheres, niosomes, dendrimers, stabilizers,
 CC buffers, excipients and additives; (7) producing (M2) a modified binding
 CC polypeptide, which is suitable for site-directed coupling, comprising
 CC modifying a polynucleotide encoding the binding polypeptide; and (8) use
 CC of a modified binding polypeptide or fusion polypeptide (III) producible
 CC by (M3) for the manufacture of a medicament or diagnostic for the
 CC prevention, treatment or diagnosis of a disease, which is characterized
 CC by an increased or decreased amount of at least one binding partner of
 CC the binding polypeptide in diseased tissue or cells involved in the
 CC disease. (I), (II) or its fusion polypeptide, or (C1) are useful for
 CC producing a medicament for the therapy of proliferative diseases, immune
 CC diseases, infectious diseases, vascular diseases, rheumatoid diseases,
 CC and diseases in which cells in or adjacent to the disease site show an

CC increased expression of EGF, or for the diagnosis of the diseases. (III)
 CC is useful for manufacturing a medicament or diagnostic for the
 CC prevention, treatment or diagnosis of the diseases. (II) is useful for
 CC site-specific coupling e.g., targeting ligands and modified human EGF and
 CC its fragments, suitable for site-specific coupling. The present sequence
 CC represents a modified mature EGF amino acid sequence, which was modified
 CC from human EGF.
 XX
 SQ Sequence 53 AA;
 Query Match 38.7%; Score 48; DB 9; Length 53;
 Best Local Similarity 53.8%; Pred. No. 9.1;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 10 DGVCYYKGVQGY 22
 ||||| : : : ||
 Db 17 DGVCMYIEALDQY 29
 RESULT 14
 ID AEB86419 standard; protein; 53 AA.
 XX
 AC AEB86419;
 XX
 DT 06-OCT-2005 (first entry)
 DE Mature modified epidermal growth factor SEQ ID NO:12.
 XX
 XX protein engineering; epidermal growth factor; cytostatic; antimicrobial;
 XX vasootropic; antirheumatic; anticholelithic; antinflammatory;
 XX proliferative disease; cancer; immune disorder; infectious disease;
 XX vascular disease.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS
 XX
 PN EP1557429-A1.
 XX
 PD 27-JUL-2005.
 XX
 PF 23-JAN-2004; 2004EP-00001454.
 XX
 PR 23-JAN-2004; 2004EP-00001454.
 XX
 PA (VECT-) VECTRON THERAPEUTICS AG.
 XX
 PI Kontermann R;
 XX
 DR WPI; 2005-573617/59.
 DR N-PsDB; AEB86434.
 XX
 PT New polynucleotide encoding a mature modified epidermal growth factor,
 PT for producing a medicament for therapy and for diagnosing e.g.
 PT proliferative diseases, immune diseases, and vascular diseases.
 XX
 PS Claim 1; SEQ ID NO 12; 50pp; English.
 XX
 CC The invention relates to a polynucleotide (I) chosen from a
 CC polynucleotides encoding a mature modified epidermal growth factor (EGF)
 CC having a 53 amino acid sequence of AEB86408 to AEB86422, where the
 CC polynucleotide has a 159 base pair sequence of AEB86423 to AEB86437, and
 CC encodes mature modified EGF. Also described: (1) a vector (VI) containing
 CC (I); (2) a host cell (H1) genetically engineered with (I) or (VI); (3)
 CC producing a modified EGF encoded by (I), comprising culturing (H1) and
 CC recovering the modified EGF encoded by (I); (4) producing (M1) cells
 CC capable of expressing modified EGF, comprising genetically engineering
 CC cells in vitro with (VI), where the modified EGF is encoded by (I); (5) a
 CC obtainable EGF (II) having the amino acid sequence encoded by (I) or
 CC (M1); (6) a composition (C1) comprising one or more (II) or
 CC a fusion polypeptide and one or more further component chosen from
 CC liposomes, virosomes, microspheres, mosomes, dendrimers, stabilizers,
 CC buffers, excipients and additives; (7) producing (M2) a modified binding

CC polypeptide, which is suitable for site-directed coupling, comprising
 CC modifying a polynucleotide encoding the binding polypeptide; and (8) use
 CC of a modified binding polypeptide or fusion polypeptide (III) producible
 CC by (M3) for the manufacture of a medicament or diagnostic for the
 CC prevention, treatment or diagnosis of a disease, which is characterized
 CC by an increased or decreased amount of at least one binding partner of
 CC the binding polypeptide in diseased tissue or cells involved in the
 CC disease. (II) (III) or its fusion polypeptide, or (C1) are useful for
 CC producing a medicament for the therapy of proliferative diseases, immune
 CC diseases, infectious diseases, vascular diseases, rheumatoid diseases,
 CC and diseases in which cells in or adjacent to the disease site show an
 CC increased expression of EGF, or for the diagnosis of the diseases. (III)
 CC is useful for manufacturing a medicament or diagnostic for the
 CC prevention, treatment or diagnosis of the diseases. (II) is useful for
 CC site-specific coupling e.g., targeting ligands and modified human EGF and
 CC its fragments, suitable for site-specific coupling. The present sequence
 CC represents a modified mature EGF amino acid sequence, which was modified
 CC from human EGF.
 XX
 SQ Sequence 53 AA;
 Query Match 38.7%; Score 48; DB 9; Length 53;
 Best Local Similarity 53.8%; Pred. No. 9.1;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 10 DGVCYYKGVQGY 22
 ||||| : : : ||
 Db 17 DGVCMYIEALDQY 29
 RESULT 15
 ID AEB86449 standard; protein; 52 AA.
 XX
 AC AEB86449;
 XX
 DT 06-OCT-2005 (first entry)
 DE Dog epidermal growth factor amino acid sequence.
 XX
 XX
 XX protein engineering; epidermal growth factor; cytostatic; antimicrobial;
 XX vasootropic; antirheumatic; anticholelithic; antinflammatory;
 XX proliferative disease; cancer; immune disorder; infectious disease;
 XX vascular disease.
 XX
 OS Canis sp.
 OS
 XX
 PN EP1557429-A1.
 XX
 PD 27-JUL-2005.
 XX
 PF 23-JAN-2004; 2004EP-00001454.
 XX
 PR 23-JAN-2004; 2004EP-00001454.
 XX
 PA (VECT-) VECTRON THERAPEUTICS AG.
 XX
 PI Kontermann R;
 XX
 DR WPI; 2005-573617/59.
 DR N-PsDB; AEB86434.
 XX
 PT New polynucleotide encoding a mature modified epidermal growth factor,
 PT for producing a medicament for therapy and for diagnosing e.g.
 PT proliferative diseases, immune diseases, and vascular diseases.
 XX
 PS Example 1; Fig 3; 50pp; English.
 XX
 CC The invention relates to a polynucleotide (I) chosen from a
 CC polynucleotides encoding a mature modified epidermal growth factor (EGF)
 CC having a 53 amino acid sequence of AEB86408 to AEB86422, where the
 CC polynucleotide has a 159 base pair sequence of AEB86423 to AEB86437, and
 CC encodes mature modified EGF. Also described: (1) a vector (VI) containing
 CC (I); (2) a host cell (H1) genetically engineered with (I) or (VI); (3)

This Page Blank (unused)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:07:19 ; Search time 9.1 Seconds

(without alignments)
243.185 Million cell updates/sec

Title: US-10-501-838a-10

Sequence: 1 TAFDFNKMIDGVCCTVYKGVQYTL 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	84	2 C86579	CT666 hypothetical
2	124	100.0	84	2 F72046	conserved hypochet
3	101	81.5	83	2 F71485	hypothetical prote
4	82	66.1	82	2 A81749	conserved hypochet
5	49	39.5	504	2 C64399	hypothetical prote
6	47	37.9	341	2 S38245	hypothetical prote
7	47	37.9	505	2 C46642	DNA primase (EC 2.
8	47	37.9	1958	2 T39808	hypothetical prote
9	46	37.1	362	2 E69605	spore coat protein
10	46	36.3	538	2 D72213	conserved hypochet
11	45	36.3	413	2 T04520	hypothetical prote
12	45	36.3	423	2 D89949	folylipolyglutamate
13	45	36.3	556	2 B64482	adenine deaminase
14	45	36.3	933	2 T43201	MutS protein homol
15	45	36.3	1116	2 S63387	probable membrane
16	45	36.3	1139	2 T20660	hypothetical prote
17	45	36.3	1466	2 T39557	vacuolar protein 8
18	45	36.3	1549	2 S50705	hypothetical prote
19	45	36.3	3305	2 T18358	apolipoprotein prec
20	44	35.5	219	2 A64340	hypothetical prote
21	44	35.5	289	2 G70449	lytB protein - Aqu
22	44	35.5	1207	2 E8HU	epidermal growth f
23	44	35.5	1579	2 S25329	carboxypeptidase Y
24	43.5	35.1	373	2 T45043	hypothetical prote
25	43.5	35.1	401	2 D83363	conserved hypochet
26	43	34.7	247	2 H69374	NH(3)-dependent NA
27	43	34.7	275	2 T17966	ribonuclease III h
28	43	34.7	367	2 W2WD5	E2 protein - human
29	43	34.7	367	2 S36524	E2 protein - human

30	43	34.7	423	2 B64079	probable amidohydr
31	43	34.7	437	2 C64113	tetrahydrofolylpol
32	43	34.7	514	2 B96653	hypothetical prote
33	43	34.7	608	2 A64992	sensor protein Aro
34	43	34.7	882	2 T39789	aminopeptidase - E
35	43	34.7	1133	1 EGRT	epidermal growth f
36	42.5	34.3	358	2 D81345	probable periplasm
37	42	33.9	16	2 PC4416	Erbb kinase activa
38	42	33.9	57	2 PC4415	Erbb kinase activa
39	42	33.9	125	2 D69940	hypothetical prote
40	42	33.9	226	2 T38234	hypothetical prote
41	42	33.9	354	2 D31751	protein kinase cat
42	42	33.9	376	2 E31751	protein kinase cat
43	42	33.9	391	2 F88778	protein T20D3.3 [I
44	42	33.9	401	2 T25031	hypothetical prote
45	42	33.9	445	2 A86767	hypothetical prote

ALIGNMENTS

RESULT 1

C86579 CT666 hypothetical protein [imported] - Chlamydia pneumoniae (strain J138)

C/Spectrum: Chlamydia pneumoniae, Chlamydia pneumoniae

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 09-Jul-2004

C/Accession: C86579

R/Shital, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Onchi, K.; Shiba, T.; I

Nucleic Acids Res. 28, 2311-2314, 2000

A/Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.

A/Reference number: A86491; MUID: 20330349; PMID: 10871362

A/Accession: C86579

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-84 <STO>

A/Cross-references: UNIPROT:Q9Z7J5; UNIPARC:UP10000139C18; GB:BA000008; NID:98979082; P

A/Experimental source: strain J138

A/Genetics:

A/Genes: CPJ0710

Query Match 100.0%; Score 124; DB 2; Length 84;
Best Match Similarity 100.0%; Pred. No. 4, 1e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAFDFNKMIDGVCCTVYKGVQYTL 23
Db 8 TAFDFNKMIDGVCCTVYKGVQYTL 30

RESULT 2

F72046 conserved hypothetical protein CP0036 [imported] - Chlamydia pneumoniae (strains CW

N/Alternate names: ct666 hypothetical protein

C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #ext_change 09-Jul-2004

C/Accession: F72046; C81620

R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.

Nature Genet. 21, 385-389, 1999

A/Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.

A/Reference number: A72000; MUID: 99206606; PMID: 10192388

A/Accession: F72046

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-84 <ARN>

A/Cross-references: UNIPROT:Q9Z7J5; UNIPARC:UP10000139C18; GB:AE001652; GB:AE001363; NI

A/Experimental source: strain CWL029

R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,

, C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.

A/Reference number: A81500; MUID: 20150255; PMID: 10684935

A/Accession: C81620

A/Status: preliminary

A:Molecule type: DNA
A:Residues: 1-84 <REA>
A:Cross-references: UNIPARC:UPI0000139C18; GB:AE002167; GB:AE002161; NID:g7188971; PIDN:
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP00710; CP0036

Query Match 100.0%; Score 124; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.1e-18;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCYVKGVOQYL 23
|||
Db 8 TAFDFNKMLDGVCYVKGVOQYL 30

RESULT 3

F71485
hypothetical protein CT666 - Chlamydia trachomatis (serotype D, strain UM3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: F71485
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac
A:Reference number: A71570; MUID:9900809; PMID:9784136
A:Accession: F71485
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <ARN>
A:Cross-references: UNIPROT:O84673; UNIPARC:UPI0000139BA6; GB:AE001337; GB:AE001273; NID:
A:Experimental source: serotype D, strain UM-3/Cx
C:Genetics:
A:Gene: CT666

Query Match 81.5%; Score 101; DB 2; Length 83;
Best Local Similarity 78.3%; Pred. No. 1.5e-08;
Matches 18; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCYVKGVOQYL 23
|||
Db 7 SAFDFNKMLDGVCYVKGVOQYL 29

RESULT 4

A81749
conserved hypothetical protein TC0037 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MOpn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: A81749
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MOpn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: A81749
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-82 <TET>
A:Cross-references: UNIPROT:G9PLQ8; UNIPARC:UPI000005778B; GB:AE002271; GB:AE002160; NID:
A:Experimental source: strain Nigg (MOpn)
C:Genetics:
A:Gene: TC0037

Query Match 66.1%; Score 82; DB 2; Length 82;
Best Local Similarity 65.2%; Pred. No. 1.3e-05;
Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCYVKGVOQYL 23
|||
Db 6 SALDFNKMLDGVCYVKGVOQYL 28

RESULT 5

CE4399
hypothetical protein MJ0795 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: CE4399
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodet, A.;
raon, J.D.; Sadow, P.M.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:9637999; PMID:8688087
A:Accession: CE4399
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-504 <BDU>
A:Cross-references: UNIPROT:O58205; UNIPARC:UPI0000139CD7; GB:U67524; GB:L77117; NID:928
A:Map position: FOR717490-719004

Query Match 39.5%; Score 49; DB 2; Length 504;
Best Local Similarity 42.1%; Pred. No. 11;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 4 DENFMKLDGVCYVKGVOQY 22
|||
Db 104 NYDKYKTYCNYTKGNQOY 122

RESULT 6

S38245
hypothetical protein - Coxiella burnetii
C:Species: Coxiella burnetii
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S38245
R:Thiele, D.; Willems, H.; Haas, M.; Kraus, H.
submitted to the EMBL Data Library, October 1993
A:Reference number: S38215
A:Accession: S38245
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-341 <THI>
A:Cross-references: UNIPROT:O45952; UNIPARC:UPI00000B4CDD; EMBL:X75356; NID:g407370; PID

Query Match 37.9%; Score 47; DB 2; Length 341;
Best Local Similarity 35.0%; Pred. No. 16;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 3 FDPFNKLDGVCYVKGVOQY 22
|||
Db 180 YDFSPDLALQTYVNVKRW 199

RESULT 7

CA6642
DNA primase (EC 2.7.7.-) 54K chain - mouse
N:Alternate names: DNA polymerase alpha/DNA primase complex 54K chain
C:Species: Mus musculus (house mouse)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: CA6642
R:Miyaazawa, H.; Izumi, M.; Tada, S.; Takada, R.; Maatani, M.; Uti, M.; Hanaoka, F.
J. Biol. Chem. 268, 8111-8122, 1993
A:Title: Molecular cloning of the cDNAs for the four subunits of mouse DNA polymerase α
A:Reference number: A6642; MUID:95216788; PMID:8463324
A:Accession: CA6642
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-505 <MTY>
A:Cross-references: UNIPROT:P33610; UNIPARC:UPI0000001986; GB:D13545; NID:g303662; PIDN:
A:Experimental source: FM3A cells
A>Note: sequence extracted from NCBI backbone (NCBI:129150, NCBI:P129151)
R:Stadlbauer, F.; Bruckner, A.; Rehfuss, C.; Eckerskorn, C.; Lottspeich, F.; Foerster,

Bur. J. Biochem. 222, 781-793, 1994

A/Title: DNA replication in vitro by recombinant DNA-polymerase-alpha-primase.

A/Reference number: S45628; MUID:94298618; PMID:8026492

A/Accession: S45629

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-108, 'G', 110-505 <STA>

A/Cross-references: UNIPARC:UPI000016CPC6; EMBL:DJ7385; NID:G9397830; PIDN:BA04203.1; PI

C/Superfamily: mouse DNA primase 54k chain

C/Keywords: DNA binding; nucleotidyltransferase

Query Match 37.9%; Score 47; DB 2; Length 505;

Best Local Similarity 44.0%; Pred. No. 23;

Matches 11; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

2 ADFPKMLDGVCTYKGVQY 22

46 ADFRVKLKALLENLGVSVYKGTQY 70

RESULT 8

hypothetical protein SPBC19C7.02 - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: T39808

R/Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Galliard, C.

submitted to the EMBL Data Library, June 1998

A/Reference number: Z21881

A/Accession: T39808

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1958 <LVN>

A/Cross-references: UNIPROT:O60152; UNIPARC:UPI0000137A05; EMBL:AL023859; PIDN:CA119570.

A/Experimental source: strain 972n-; cosmid c19C7

C/Genetics:

A/Status: SPDB:SPBC19C7.02

A/Map position: 2

Query Match 37.9%; Score 47; DB 2; Length 1958;

Best Local Similarity 35.0%; Pred. No. 93;

Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

4 DFKMLDGVCTYKGVQY 23

599 DFPNLFQMGCPYTRAVSQH 618

RESULT 9

spore coat protein (inner) coth - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C/Accession: E69605

R/Kuner, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berti

C./ Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A./ Ehrlich, S.D.; Emerson, P.T.; Eutani, K.D.; Ewington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallazzi, A.; Gall

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holappell, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsfeld, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidis, A.; Lardinois,

A/Authors: Lander, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue

Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portecelle

Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A/Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Ser

akeuchi, M.; Yamakoshi, A.; Tanaka, T.; Terpest, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T./ Wiltner, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A/Authors: Yoshikawa, H.F.; Zumbstein, B.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: E69605

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-362 <KUN>

A/Cross-references: UNIPROT:Q45535; UNIPARC:UPI0000060B5A; GB:Z99122; GB:AL009126; NID:9

A/Experimental source: strain 168

C/Genetics:

A/Status: coth

Query Match 37.1%; Score 46; DB 2; Length 362;

Best Local Similarity 42.1%; Pred. No. 24;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

5 FPKMLDGVCTYKGVQY 23

337 FDRPDDVCYIKKNSQYL 355

RESULT 10

conserved hypothetical protein - Thermotoga maritima (strain MSB8)

C/Species: Thermotoga maritima

C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C/Accession: D72213

R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hicke

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.

C.M.

Nature 399, 323-329, 1999

A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se

A/Reference number: A72200; MUID:99287316; PMID:10360571

A/Accession: D72213

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-538 <ARN>

A/Cross-references: UNIPROT:Q9X294; UNIPARC:UPI00000C1209; GB:AE001815; GB:AE00512; NII

A/Experimental source: strain MSB8

C/Genetics:

A/Status: TM1773

Query Match 37.1%; Score 46; DB 2; Length 538;

Best Local Similarity 36.8%; Pred. No. 35;

Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

3 FPKMLDGVCTYKGVQ 21

497 FDRRLSDVIAFLQSIQ 515

RESULT 11

hypothetical protein F16A16.100 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004

C/Accession: T04520

R/Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Hohnsbeil, J.; Mew

submitted to the Protein Sequence Database, February 1999

A/Reference number: Z15376

A/Accession: T04520

A/Molecule type: DNA

A/Residues: 1-413 <BEV>

A/Cross-references: UNIPROT:Q9SVU6; UNIPARC:UPI00000A6306; EMBL:AL035353

A/Experimental source: cultivar Columbia; BAC clone F16A16

C/Genetics:

A/Map position: 4

A/Intons: 22/2; 262/3; 287/3; 309/3; 331/3

A/Note: F16A16.100

Query Match 36.3%; Score 45; DB 2; Length 413;

Best Local Similarity 43.5%; Pred. No. 39;

Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

1 TAFDFKMLDGVCTYKGVQY 23

308 TKTRSSMLDDVIEVYLSQSI 330

```
RESULT 12
D89949
A:Accession: T43201
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-933 <WLN>
A:Cross-references: UNIPROT:Q19272; UNIPARC:UPI000016B67D; EMBL:AF070070; NID:G3831700;
C:Genetics:
A:Gene: mbh-5

Query Match          36.3% Score 45; DB 2; Length 933;
Best Local Similarity 42.1% Pred. No. 89;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Cy 3 FDFKMLDGVCTYKGVQ 21
Db 683 FTVDSVLDGMSYFAKDVQ 701

RESULT 15
S63397
A:Accession: YNR065C - yeast (Saccharomyces cerevisiae)
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1116 <DUE>
A:Cross-references: UNIPROT:P53751; UNIPARC:UPI000013BAC9; EMBL:E71680; NID:G1302593; PI
F:990-1006/Domain: transmembrane #status predicted <TM1>
F:990-1006/Domain: transmembrane #status predicted <TM2>

Query Match          36.3% Score 45; DB 2; Length 1116;
Best Local Similarity 47.1% Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Cy 3 FDFKMLDGVCTYKGV 19
Db 858 YNFYKASDGTCKLVKGL 874

Search completed: January 23, 2006, 09:32:54
Job time : 10.1 secs
```

```
RESULT 13
B64482
A:Accession: B64482
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-556 <BLU>
A:Cross-references: UNIPROT:Q58854; UNIPARC:UPI0000125518; GB:U67586; GB:L77117; NID:G15
C:Genetics:
A:Map position: FOR1428283-1429953
A:Superfamily: adenine deaminase adeC
C:Keywords: hydrolase

Query Match          36.3% Score 45; DB 2; Length 556;
Best Local Similarity 50.0% Pred. No. 52;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Cy 5 FFKMLDGVCTYKKG 18
Db 183 YNKLDIGHCPRKKG 196

RESULT 14
T43201
A:Accession: T43201
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-933 <WLN>
A:Cross-references: UNIPROT:Q19272; UNIPARC:UPI000016B67D; EMBL:AF070070; NID:G3831700;
C:Genetics:
A:Gene: mbh-5

Query Match          36.3% Score 45; DB 2; Length 933;
Best Local Similarity 42.1% Pred. No. 89;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Cy 3 FDFKMLDGVCTYKGVQ 21
Db 683 FTVDSVLDGMSYFAKDVQ 701

RESULT 15
S63397
A:Accession: YNR065C - yeast (Saccharomyces cerevisiae)
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1116 <DUE>
A:Cross-references: UNIPROT:P53751; UNIPARC:UPI000013BAC9; EMBL:E71680; NID:G1302593; PI
F:990-1006/Domain: transmembrane #status predicted <TM1>
F:990-1006/Domain: transmembrane #status predicted <TM2>

Query Match          36.3% Score 45; DB 2; Length 1116;
Best Local Similarity 47.1% Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Cy 3 FDFKMLDGVCTYKGV 19
Db 858 YNFYKASDGTCKLVKGL 874

Search completed: January 23, 2006, 09:32:54
Job time : 10.1 secs
```

```
RESULT 12
D89949
A:Accession: B89949
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-423 <KUR>
A:Cross-references: UNIPROT:Q99TJ9; UNIPARC:UPI0000054C15; GB:BA000018; PID:G13701459; F
A:Experimental source: strain N315
C:Genetics:
A:Gene: folC
A:Superfamily: folY1polyglutamate synthase

Query Match          36.3% Score 45; DB 2; Length 423;
Best Local Similarity 77.8% Pred. No. 40;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 4 DFNKMLDGV 12
Db 280 DFNKMLDGI 288

RESULT 13
B64482
A:Accession: B64482
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-556 <BLU>
A:Cross-references: UNIPROT:Q58854; UNIPARC:UPI0000125518; GB:U67586; GB:L77117; NID:G15
C:Genetics:
A:Map position: FOR1428283-1429953
A:Superfamily: adenine deaminase adeC
C:Keywords: hydrolase

Query Match          36.3% Score 45; DB 2; Length 556;
Best Local Similarity 50.0% Pred. No. 52;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Cy 5 FFKMLDGVCTYKKG 18
Db 183 YNKLDIGHCPRKKG 196

RESULT 14
T43201
A:Accession: T43201
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-933 <WLN>
A:Cross-references: UNIPROT:Q19272; UNIPARC:UPI000016B67D; EMBL:AF070070; NID:G3831700;
C:Genetics:
A:Gene: mbh-5

Query Match          36.3% Score 45; DB 2; Length 933;
Best Local Similarity 42.1% Pred. No. 89;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Cy 3 FDFKMLDGVCTYKGVQ 21
Db 683 FTVDSVLDGMSYFAKDVQ 701

RESULT 15
S63397
A:Accession: YNR065C - yeast (Saccharomyces cerevisiae)
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1116 <DUE>
A:Cross-references: UNIPROT:P53751; UNIPARC:UPI000013BAC9; EMBL:E71680; NID:G1302593; PI
F:990-1006/Domain: transmembrane #status predicted <TM1>
F:990-1006/Domain: transmembrane #status predicted <TM2>

Query Match          36.3% Score 45; DB 2; Length 1116;
Best Local Similarity 47.1% Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Cy 3 FDFKMLDGVCTYKGV 19
Db 858 YNFYKASDGTCKLVKGL 874

Search completed: January 23, 2006, 09:32:54
Job time : 10.1 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:05:19 ; Search time 60.7 Seconds
(without alignments)
267,334 Million cell updates/sec

Title: US-10-501-838a-10
Perfect score: 124
Sequence: 1 TAFDFKMLDGVCTYKGVQYL 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_05.80: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	84	1 Y710 CHLPN	Q92715 chlamydia p
2	101	81.5	83	1 Y666 CHLTR	Q84673 chlamydia t
3	94	75.8	84	2 O51777 CHLAB	Q51777 chlamydia t
4	94	75.8	86	2 O824V7 CHLCV	Q824V7 chlamydia t
5	82	66.1	82	1 Y037 CHLMU	Q9148 chlamydia m
6	52.5	42.3	500	2 Q4XYW8 PLACH	Q4XYW8 plasmidium
7	50	40.3	818	2 Q7PDM0 PLAYO	Q7PDM0 plasmidium
8	50	40.3	1852	2 Q812U9 PLAF7	Q812U9 plasmidium
9	49	39.5	504	1 Y795 MBRJA	Q58205 methanococc
10	49	39.5	915	2 Q4PR49 USTMA	Q4PR49 uetiliago ma
11	48.5	39.1	394	1 MFNA PTRKO	Q51362 pyrococcus
12	48	38.7	118	2 Q4G213 YVIRU	Q4G213 ludwigia ye
13	48	38.7	293	2 Q69205 9CLOS	Q69205 citrus tris
14	48	38.7	293	2 Q69206 9CLOS	Q69206 citrus tris
15	48	38.7	293	2 Q69207 9CLOS	Q69207 citrus tris
16	48	38.7	293	2 Q69208 9CLOS	Q69208 citrus tris
17	48	38.7	293	2 Q69209 9CLOS	Q69209 citrus tris
18	48	38.7	293	2 Q69210 9CLOS	Q69210 citrus tris
19	48	38.7	348	2 Q4YX26 PLABE	Q4YX26 plasmidium
20	48	38.7	516	1 NACH DROAN	Q61370 dirosophila
21	48	38.7	557	2 Q8XOK3 RALSO	Q8XOK3 raleconia s
22	48	38.7	719	1 TIP PLAF7	Q81317 plasmidium
23	48	38.7	791	2 Q94175 PNECA	Q94175 pneumocysti
24	48	38.7	940	2 Q94176 PNECA	Q94176 pneumocysti
25	48	38.7	1024	2 Q6AHU3 PNECA	Q6AHU3 pneumocysti
26	48	38.7	1564	2 Q4N477 THEPA	Q4N477 thelaxia p
27	48	38.7	3107	2 P87587 9CLOS	P87587 citrus tris
28	48	38.7	3115	2 Q10467 9CLOS	Q10467 citrus tris
29	48	38.7	3115	2 Q91FX0 9CLOS	Q91FX0 citrus tris
30	48	38.7	3115	2 Q9WID7 9CLOS	Q9WID7 citrus tris
31	48	38.7	3132	2 Q9DTG5 9CLOS	Q9DTG5 citrus tris

32	47	37.9	185	2 Q4F953 IPOBA	Q4F953 ipomoea bat
33	47	37.9	309	2 O52888 COXBU	O52888 coxiella bu
34	47	37.9	309	2 Q956J4 COXBU	Q956J4 coxiella bu
35	47	37.9	313	2 Q45910 COXBU	Q45910 coxiella bu
36	47	37.9	341	2 Q45952 COXBU	Q45952 coxiella bu
37	47	37.9	484	2 Q4VBC0 RAT	Q4VBC0 rattus norv
38	47	37.9	505	1 PR12_MOUSE	P33610 mus musculu
39	47	37.9	507	1 PR12_MOUSE	O89044 rattus norv
40	47	37.9	510	2 Q741DE LACJO	Q741DE lactobacill
41	47	37.9	535	1 NACH DROME	O61365 dirosophila
42	47	37.9	1213	2 O55EN6 DICDI	O55EN6 dictyostell
43	47	37.9	1958	1 UBR1 SCHPO	O60152 schizosacch
44	46.5	37.5	715	2 UBR1 SCHPO	O7N6M4 photorhabd
45	46	37.1	207	2 Q4ZGM3 YVIRU	Q4ZGM3 macrobrachi

ALIGNMENTS

RESULT 1
ID Y710 CHLPN STANDARD; PRT; 84 AA.
AC Q92715;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 47, Last annotation update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Protein CPN0710/CP0036/CPJ0710/CPB0737.
GN Ordered locus names=CPN0710, CP0036, CPJ0710, CPB0737;
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_Taxid=83558;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CM1029;
RX MEDLINE=9206606; PubMed=10192388; DOI=10.1038/7716;
RA Kalman S., Mitchell W.P., Marache R., Lamme C.J., Fan J., Hyman R.W.,
Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
White O., Hickey B.K., Peterson J.D., Uetrich T.R., Berry K.J.,
Bass S., Linher K.D., Weidman J.P., Knout H.M., Craven B., Bowman C.,
Dodson R.J., Gilm M.L., Nelson W.C., Deboy R.T., Kolonay J.F.,
McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=J138;
RX MEDLINE=2030339; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CML029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=J138;
RA Hirakawa H., Shirai M., Kuhara S.;
RT "Genomic sequence comparison of two unrelated isolates of Chlamydia
pneumoniae from Japan and U.S.";
RN Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TW-183;
RA Geng M.W., Schumacher A., Muehldorfer I., Benesch K.W., Schaefer K.P.,
Schneider S., Pohl T., Esig A., Martre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with

RT other Chlamydia strains based on whole genome sequence analysis.;"

RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the Chlamydial CPN0710/CT666/TC0037 family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

DR EMBL; AE001652; AAD18849.1; -; Genomic DNA.

DR EMBL; AE002167; AAF37931.1; -; Genomic DNA.

DR EMBL; BA000008; BAA98917.1; -; Genomic DNA.

DR EMBL; AB035952; BAA88660.1; -; Genomic DNA.

DR EMBL; AE017153; AAP88666.1; -; Genomic DNA.

DR PIR; C86579; C86579.

DR PIR; F72046; F72046.

DR PHCI-2DPAGE; Q92705; -.

DR TIGR; CP0036; -.

KW Complete proteome.

SO SEQUENCE 84 AA; 9348 MW; 50A9CC271A5CB693 CRC64;

Query Match 100.0%; Score 124; DB 1; Length 84;

Best Local Similarity 100.0%; Pred. No. 1.7e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAFDFNKMULDGVCTYKGVQOYL 23

DB 8 TAFDFNKMULDGVCTYKGVQOYL 30

RESULT 2

Y666 CHLTR STANDARD; PRT; 83 AA.

AC 084673;

DT 16-OCT-2001 (Rel. 40; Created)

DT 16-OCT-2001 (Rel. 40; Last sequence update)

DT 10-MAY-2005 (Rel. 47; Last annotation update)

DE Protein CT666.

GN OrderedLocusNames=CT666;

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiae; Chlamydiaceae; Chlamydia.

OX NCBI_TaxId=813;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=D/WR-3/Cx;

RX MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;

RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;

RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis."

RT Science 282:754-759(1998).

RL -1- SIMILARITY: Belongs to the Chlamydial CPN0710/CT666/TC0037 family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

DR EMBL; AE001337; AAC68261.1; -; Genomic DNA.

DR PIR; F71485; F71485.

DR PHCI-2DPAGE; O84673; -.

KW Complete proteome.

SO SEQUENCE 83 AA; 9160 MW; FF227A1F3549AAE3 CRC64;

Query Match 81.5%; Score 101; DB 1; Length 83;

Best Local Similarity 78.3%; Pred. No. 7e-08;

Matches 18; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 TAFDFNKMULDGVCTYKGVQOYL 23

DB 8 TAFDFNKMULDGVCTYKGVQOYL 30

DB 7 SAFNFNKMULDGVCKYGVQOYL 29

RESULT 3

O5L777 CHLAB PRELIMINARY; PRT; 84 AA.

AC O5L777

DT 01-FEB-2005 (TRENBLREL. 29; Created)

DT 01-FEB-2005 (TRENBLREL. 29; Last sequence update)

DT 01-FEB-2005 (TRENBLREL. 29; Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=CHAB033;

OS Chlamydia abortus.

OC Bacteria; Chlamydiae; Chlamydiaceae; Chlamydia.

OX NCBI_TaxId=83555;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=S26/3;

RX PubMed=15837807; DOI=10.1101/gr.3684805;

RA Thomson N.R., Yeats C., Bell K., Holden M.T.G., Bentley S.D., Livingston M., Cerceno-Tarraga A.-M., Harris B., Doggett J., Ormond D., Mungall K., Clarke K., Felwell T., Hance Z., Sanders M., Quail M.A., Price C., Barrell B.G., Parkhill J., Longbottom D.;

RA "The Chlamydia abortus genome sequence reveals an array of RT variable proteins that contribute to interspecies variation.;"

RL Genome Res. 15:629-640(2005).

DR EMBL; CR848038; CAH63491.1; -; Genomic DNA.

KW Complete proteome; Hypothetical protein.

SO SEQUENCE 84 AA; 9234 MW; E36BA35275A6949 CRC64;

Query Match 75.8%; Score 94; DB 2; Length 84;

Best Local Similarity 73.9%; Pred. No. 8.9e-07;

Matches 17; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 TAFDFNKMULDGVCTYKGVQOYL 23

DB 8 SAFNFNKMULDGVCKYGVQOYL 30

RESULT 4

O824V7 CHLCV PRELIMINARY; PRT; 86 AA.

AC O824V7;

DT 01-JUN-2003 (TRENBLREL. 24; Created)

DT 01-JUN-2003 (TRENBLREL. 24; Last sequence update)

DT 01-JUN-2003 (TRENBLREL. 24; Last annotation update)

DE Hypothetical protein.

GN OrderedLocusNames=CCA00032;

OS Chlamydia caviae.

OC Bacteria; Chlamydiae; Chlamydiaceae; Chlamydia.

OX NCBI_TaxId=83557;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=GPIC;

RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gk9321;

RA Read T.D., Myers G.S.A., Brumham R.C., Nelson W.C., Paulsen I.T., Heidelberg J.F., Holtzapfe E.K., Khouri H.M., Fedorova N.B., Carey H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J., White O., Salzberg S.L., Hefi R.-C., McClarty G., Rank R.G., Bayoil P.M., Fraser C.M.;

RA "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC): RT examining the role of niche-specific genes in the evolution of the Chlamydiaceae.;"

RT Nucleic Acids Res. 31:2134-2147(2003).

RL Nucleic Acids Res. 31:2134-2147(2003).

DR EMBL; AE016994; AAP04784.1; -; Genomic DNA.

DR TIGR; CCA00032; -.

KW Complete proteome; Hypothetical protein.

SO SEQUENCE 86 AA; 9378 MW; 7F5E937FFA86478B CRC64;

Query Match 75.8%; Score 94; DB 2; Length 86;

Best Local Similarity 73.9%; Pred. No. 9.1e-07;

Matches 17; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TAFDPMKLDGVCYTVKGVQOYL 23
 ID Y037 CHLMU STANDARD; PRT; 82 AA.
 DB 10 SAFNFMNDMLNGVCYKVVGVQOYL 32

RESULT 5
 Y037 CHLMU STANDARD; PRT; 82 AA.

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypochemical protein TC0037.
 GN OrderedlocusNames=TC0037;
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OK NCBI_TaxId=83560;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RP STRAIN=MOPn / N199;
 RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 White O., Hickey K.K., Peterson J.D., Utterback T.R., Berry K.J.,
 Bass S., Linher K.D., Weidman J.F., Knout H.M., Craven B., Bowman C.,
 Dodson R.J., Gwin M.L., Nelson W.C., Deboy R.T., Kolonay J.F.,
 McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
 RA "Genome sequences of Chlamydia trachomatis MOPn and Chlamydia
 pneumoniae AR39".

RT Nucleic Acids Res. 28:1397-1406(2000).
 CC -1- SIMILARITY: Belongs to the chlamydial Cpn0710/CT666/TC0037 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC EMBL; AE002271; AAP38928.1; -; Genomic_DNA.
 DR PIR; A81749; A81749.
 DR TIGR; TC0037; -

KW Complete proteome; Hypochemical protein.
 SQ SEQUENCE 82 AA; 9069 MW; 28F7C6BD914F6345 CRC64;

Query Match 66.1%; Score 82; DB 1; Length 82;

Best Local Similarity 65.2%; Pred. No. 6,7e-05;

Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TAFDPMKLDGVCYTVKGVQOYL 23
 ID Y037 CHLMU STANDARD; PRT; 82 AA.
 DB 10 SAFNFMNDMLNGVCYKVVGVQOYL 32

RESULT 6
 Q4XYW8 PLACH PRELIMINARY; PRT; 500 AA.

DT 13-SEP-2005 (TREMBlrel. 31, Created)

DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)

DE Hypochemical protein (fragment).

GN ORFNames=PC000598.02.0;
 OS Plasmodium chabaudi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OK NCBI_TaxId=5825;
 RN NUCLEOTIDE SEQUENCE.

RP Hall N., Karrae M., Raine J.D., Carlton J.M., Kool J.T.W.A.,
 RA Bertram M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
 RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
 RA Quail M.A., Ormond D., Doggett J., Trueman H.B., Mendoza J.,
 RA Bidwell S.L., Rajadwaram M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
 RA Jansz C.J., Barrett B., Turner C.M., Waters A.P., Sinden R.S.;
 RT "A comprehensive survey of the Plasmodium life cycle by genomic,

RT transcriptomic, and proteomic analyses.";

CC Science 307:82-86(2005).

CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; CAJ01002196; CAH77892.1; -; Genomic_DNA.
 DR InterPro; IPR012287; Homeodomain-rel.
 DR InterPro; IPR001005; Myb_DNA_Bd.
 DR Pfam; PF00249; Myb_DNA-binding; 1.
 DR SMART; SM00717; SANT; 1.
 DR PROSITE; PSS0090; MYB 3; 1.
 KW Hypochemical protein; Nuclear protein.
 FT NON_TER

SQ SEQUENCE 500 AA; 58203 MW; 9EE90BE4C202ABE1 CRC64;

Query Match 42.3%; Score 52.5; DB 2; Length 500;
 Best Local Similarity 34.6%; Pred. No. 19;
 Matches 9; Conservative 6; Mismatches 8; Indels 3; Gaps 1;

Qy 1 TAFDPMKLDGVCYTVKGVQOYL 23
 ID Y037 CHLMU STANDARD; PRT; 82 AA.
 DB 10 SAFNFMNDMLNGVCYKVVGVQOYL 32

RESULT 7
 Q7PDM0 PLAYO PRELIMINARY; PRT; 818 AA.

DT 01-MAR-2004 (TREMBlrel. 26, Created)

DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)

DE Erythrocyte membrane protein PFEmp3.

GN Name=PY05789;

OS Plasmodium yoelii yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OK NCBI_TaxId=73239;
 RN NUCLEOTIDE SEQUENCE.

RP STRAIN=17XNL;
 RX MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;

RA Carlton J.M., Angluoi S.V., Suh B.B., Kool J.T.W., Perera M.,
 RA Silva J.C., Ermolova M.D., Allen J.B., Selengut J.D., Koo H.L.,
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
 RA Shallow S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
 RA Cho J.K., Quackenbush J., Sedegah M., Shoib A., Cummings L.M.,
 RA Florens L., Yates J.R., III, Raine J.D., Sinden R.E., Harris M.A.,
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
 RA van Lin L.H., Jansz C.J., Waters A.P., Smith H.O., White O.R.,
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
 RA Carucci D.J.;

RT "Genome sequence and comparative analysis of the model rodent malaria
 parasite Plasmodium yoelii yoelii.";

CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; AABL01001886; EAA17871.1; -; Genomic_DNA.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR011591; BclIIlum.
 DR InterPro; IPR008627; GETHR.
 DR InterPro; IPR001005; Myb_DNA-binding.
 DR Pfam; PF05671; GETHR; 3.
 DR Pfam; PF00249; Myb_DNA-binding; 1.
 DR Prodom; PD001963; BclIIlum; 1.
 DR PROSITE; PSS0090; MYB 3; 1.
 KW Nuclear protein.
 SQ SEQUENCE 818 AA; 93391 MW; 12587D693CD3F878 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 818;
 Best Local Similarity 46.7%; Pred. No. 79;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

Qy      1 TAFDFNKMLDGVCTY 15
Db      530 SAYEIRKILIDGVCKY 544

RESULT 8
081209 PLAF7 PRELIMINARY; PRT; 1852 AA.
AC 081209
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein PF11000w.
GN Name=PF11000w;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
ON NCBI_TaxId=36329;
RX NUCLEOTIDE SEQUENCE.
MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Felwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphrey S., Jagels K., James K.D., Johnson D., Kethorou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Leonard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
RL Nature 419:527-531 (2002).
DR EMBL: AL929357; CAD51866.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1852 AA; 224486 MW; D3BBBD73663951D CRC64;

Query Match
Best Local Similarity 40.3%; Score 50; DB 2; Length 1852;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      5 FRKMLDGVCTTY 16
Db      554 FRKMLDGVCTTY 565

RESULT 9
Y795_METJA STANDARD; PRT; 504 AA.
ID Y795_METJA
AC 058205;
DT 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 10-MAY-2005 (rel. 47, Last annotation update)
DE Hypothetical protein MJ0795.
GN OrderedlocusNames=MJ0795;
OC Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
ON NCBI_TaxId=2190;
RX NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RA MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weisscock K.G., Merrick J.D., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhlmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

```

```

RA Corton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073 (1996).
CC -! SIMILARITY: To M.jannaschii MJ1506 and MJ1561.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: U67524; AAB98798.1; -; Genomic_DNA.
DR PIR: C64399; C64399.
DR TIGR: MJ0795; -.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 6
SQ SEQUENCE 504 AA; 57147 MW; A1A659DB7EAE674B CRC64;

Query Match
Best Local Similarity 39.5%; Score 49; DB 1; Length 504;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy      4 DFNKMLDGVCTTYKGVQOY 22
Db      104 NYDKITGYCNYTKGNQOY 122

RESULT 10
Q4PE49 USTMA PRELIMINARY; PRT; 915 AA.
ID Q4PE49
AC Q4PE49
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM01614.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
ON NCBI_TaxId=237631;
RX NUCLEOTIDE SEQUENCE.
RA STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-Zehra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Archachhi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Biltsheteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Chehatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Dufley N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearlin G., Gnere S.,
RA Gurke A., Guyette A., Graham J., Grandols E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Homan T., Horn A., Houde N., Hughes L., Hulme W., Huby E., Iliev I.,
RA Jaffe D., Jones C., Kamel M., Kamat A., Kamysheleis M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutis K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-Toh K., Liu X., Lokytang T., Lokytang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabelli A., Maru K., Matthews C., Meneus L.,
RA McCarthy M., McDonough S., McGhee T., Meldrum J., Meneus L.,
RA Mestrov J., Mihalev A., Minova T., Mikkelson T., Mienga V., Moru K.,
RA Moses J., Multran L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizati M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkun P., Piyant B.,
RA Pucelli S., Rachupka T., Ramesam U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,

```

RA Rutman M., Schupbach R., Seaman C., Settivali S., Sharpe T.,
 RA Sheidman J., Shera N., Shi J., Smirnov S., Smith C., Sougnuez C.,
 RA Spencer B., Stalker J., Stange-Thomann N., Stavropoulos S.,
 RA Stenson K., Stone C., Stone S., Stubbs M., Talamas J., Tchinga P.,
 RA Tensing P., Teste S., Theodore J., Thouloung Y., Topham K.,
 RA Towey S., Tsamla T., Tjomo N., Vallee D., Vassiliev H.,
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
 RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
 RA Yang S., Yeager S., Yee E., Young G., Zaimoun J., Zembek L.,
 RA Zimmer A., Zody M., Zander E.,
 RT "The genome sequence of *Ustilago maydis*."
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC EMBL, AACF0100061; EAK02367.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 915 AA; 102420 MW; 5D3B897B18D97C36 CRC64;

Query Match 39.5%; Score 49; DB 2; Length 915;
 Best Local Similarity 44.4%; Pred. No. 1.3e+02;
 Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 4 DFNKMLDGVCTYVK 21
 DB 464 DFDLTDLDMSTFVKGLSE 481

RESULT 11
 ID MFNA PYRKO STANDARD; PRT; 384 AA.
 AC 05J382;
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE L-tyrosine decarboxylase (EC 4.1.1.25) (TDC).
 GN Name:mfna; OrderedlocusNames=TK1814;
 OS Pyrococcus kodakarensis (Thermococcus kodakarensis).
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Thermococcus.
 OX NCBI_TaxID=69014;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=KOD1;
 RX PubMed=15710748; DOI=10.1101/gr.3003105;
 RA Fukui T., Atomi H., Kanai T., Matsumi R., Fujiwara S., Imanaka T.;
 RT "Complete genome sequence of the hyperthermophilic archaeon
 RT Thermococcus kodakarensis KOD1 and comparison with Pyrococcus
 RT genomes.";
 RL Genome Res. 15:352-363(2005).
 CC -1- FUNCTION: Specifically catalyzes the decarboxylation of L-tyrosine
 CC to produce tyramine (By similarity).
 CC -1- CATALYTIC ACTIVITY: L-tyrosine = tyramine + CO(2).
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SIMILARITY: Belongs to the group II decarboxylase family. Archaeal
 CC L-tyrosine decarboxylase subfamily.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, AP006878; BAD86003.1; -; Genomic_DNA.
 DR HAMAP, MF_01610; -; 1.
 DR InterPro, IPR002129; Pyridoxal dec.
 DR PANTHER, PTHR11999; Pyridoxal dec; 1.
 DR Pfam, PF00282; Pyridoxal dec; 1.
 DR PROSITE, PS00392; DDC_GAD_HDC_YDC; 1.
 KW Complete proteome; Decarboxylase; Lyase; Pyridoxal phosphate.
 FT BINDING 231
 FT 231 Pyridoxal phosphate (covalent) (By
 FT similarity).

SQ SEQUENCE 384 AA; 42545 MW; 3FAF757953FA9C33 CRC64;
 Query Match 39.1%; Score 48.5; DB 1; Length 384;
 Best Local Similarity 42.3%; Pred. No. 62;
 Matches 11; Conservative 4; Mismatches 6; Indels 5; Gaps 1;

QY 3 DFNKMLDGVCTYVK 23
 DB 26 FDSGKILGSMCTYPPFAVKVMKXT 51

RESULT 12
 ID Q4GZ13_YVIRU PRELIMINARY; PRT; 118 AA.
 AC Q4GZ13;
 DT 13-SEP-2005 (TRMBLrel. 31, Created)
 DT 13-SEP-2005 (TRMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TRMBLrel. 31, Last annotation update)
 DE C1 protein.
 GN Name=c1;
 OS Ludwigia yellow vein virus-associated DNA beta.
 OC Viruses; Satellite; Satellite Nucleic Acids;
 OC Begomovirus-associated DNA beta-like sequences.
 OX NCBI_TaxID=325113;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=G37;
 RA Huang J., Zhou X., Jiang T.;
 RT "Molecular characterization of two distinct begomoviruses infecting
 RT Ludwigia hyssopifolia.";
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AJ965541; CA184288.1; -; Genomic_DNA.
 SQ SEQUENCE 118 AA; 13420 MW; 913C78P9C8706DD5 CRC64;

Query Match 38.7%; Score 48; DB 2; Length 118;
 Best Local Similarity 46.7%; Pred. No. 22;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 DFNKMLDGVCTYVK 17
 DB 57 FDFNSLEEGICNLK 71

RESULT 13
 ID 069205_9CLOS PRELIMINARY; PRT; 293 AA.
 AC 069205;
 DT 25-OCT-2004 (TRMBLrel. 28, Created)
 DT 25-OCT-2004 (TRMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TRMBLrel. 28, Last annotation update)
 DE P349 (Fragment).
 OS Citrus tristeza virus.
 OC Citrus tristeza virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
 OC Closterovirus.
 OX NCBI_TaxID=12162;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Herrera-Isidron L., Hernandez-Valencia M.C., Sanchez-Briseno V.,
 RA Iturriza-Garciduenas D.R., Ochoa-Sanchez J.C., Rivera-Bustamante R.,
 RA Martinez-Sotiano J.P.;
 RT "Molecular analysis of Citrus tristeza virus isolates from Mexico.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AY652915; AAT75285.1; -; Genomic_RNA.
 FT NON TER 1
 FT 1
 FT NON TER 293
 FT 293
 SQ SEQUENCE 293 AA; 32161 MW; 170A2DC26A948AF CRC64;

Query Match 38.7%; Score 48; DB 2; Length 293;
 Best Local Similarity 57.1%; Pred. No. 56;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 DFNKMLDGVCTYVK 17
 DB 464 DFDLTDLDMSTFVKGLSE 481

Db 113 DFDELCHAVCTYVK 126

RESULT 14

ID 069206_9CLOS PRELIMINARY; PRT; 293 AA.
 AC 069206;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE P349 (Fragment).
 OS Citrus tristeza virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
 OC Closterovirus.
 OX NCBI_TaxID=12162;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Herrera-Isidron L., Hernandez-Valencia M.C., Sanchez-Briseno V.,
 RA Iturrizaga-Garciduenas D.R., Ochoa-Sanchez J.C., Rivera-Bustamante R.,
 RA Martinez-Soriano J.P.;
 RT "Molecular analysis of Citrus tristeza virus isolates from Mexico.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY652914; AAT75284.1; -; Genomic_RNA.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 293 AA; 32000 MW; 82B2464AC4E6348 CRC64;

Query Match

Best Local Similarity 38.7%; Score 48; DB 2; Length 293;
 Best Local Similarity 57.1%; Pred. No. 56;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 DFNKMLDGVCTYVK 17
 ||::: |||||
 Db 113 DFDELCHAVCTYVK 126

RESULT 15

ID 069207_9CLOS PRELIMINARY; PRT; 293 AA.
 AC 069207;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE P349 (Fragment).
 OS Citrus tristeza virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
 OC Closterovirus.
 OX NCBI_TaxID=12162;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Herrera-Isidron L., Hernandez-Valencia M.C., Sanchez-Briseno V.,
 RA Iturrizaga-Garciduenas D.R., Ochoa-Sanchez J.C., Rivera-Bustamante R.,
 RA Martinez-Soriano J.P.;
 RT "Molecular analysis of Citrus tristeza virus isolates from Mexico.";
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY652913; AAT75283.1; -; Genomic_RNA.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 293 AA; 32066 MW; 43CCB67B0D6176A9 CRC64;

Query Match

Best Local Similarity 38.7%; Score 48; DB 2; Length 293;
 Best Local Similarity 57.1%; Pred. No. 56;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 DFNKMLDGVCTYVK 17
 ||::: |||||
 Db 113 DFDELCHAVCTYVK 126

Search completed: January 23, 2006, 09:31:17
 Job time : 62.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:07:40 ; Search time 16.6 Seconds
(without alignments)

114,551 Million cell updates/sec

Title: US-10-501-838a-10

Perfect score: 124

Sequence: 1 TAFDPMKMLDGVCTYKGVQYXL 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5/COMB.pep:*
2: /cgn2_6/prodata/1/aa/6/COMB.pep:*
3: /cgn2_6/prodata/1/aa/H/COMB.pep:*
4: /cgn2_6/prodata/1/aa/PCTUS/COMB.pep:*
5: /cgn2_6/prodata/1/aa/RE/COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	100.0	91	2	US-09-198-452A-753
2	124	100.0	91	2	US-09-438-185A-712
3	50	40.3	520	2	US-09-902-540-10545
4	48	38.7	44	1	US-08-838-957A-23
5	48	38.7	53	6	5332669-3
6	45	36.3	371	2	US-09-533-029-110
7	45	36.3	427	2	US-09-134-001C-3049
8	44	35.5	37	1	US-08-039-364-14
9	44	35.5	37	1	US-09-158-710-14
10	44	35.5	46	2	US-08-168-091A-44
11	44	35.5	46	2	US-08-889-437-12
12	44	35.5	46	2	US-09-126-121-12
13	44	35.5	47	2	US-08-753-007A-17
14	44	35.5	47	2	US-09-398-496-17
15	44	35.5	47	2	US-09-553-769-8
16	44	35.5	48	2	US-09-020-880-15
17	44	35.5	48	2	US-08-915-096A-13
18	44	35.5	48	2	US-09-101-544-15
19	44	35.5	48	6	5434135-3
20	44	35.5	50	1	US-09-097-661-12
21	44	35.5	53	1	US-07-869-176-1
22	44	35.5	53	1	US-08-284-923-1
23	44	35.5	53	1	US-08-360-841-2
24	44	35.5	53	1	US-08-861-000-1
25	44	35.5	53	1	US-08-619-032B-1
26	44	35.5	53	2	US-09-530-685A-4
27	44	35.5	53	2	US-09-848-664A-31

28	44	35.5	53	6	5434135-2	Patent No. 5434135
29	44	35.5	53	6	5218093-1	Patent No. 5218093
30	44	35.5	58	2	US-09-249-697A-11	Sequence 11, Appl
31	44	35.5	58	2	US-09-363-316B-11	Sequence 11, Appl
32	44	35.5	58	2	US-10-136-227A-11	Sequence 11, Appl
33	44	35.5	58	2	US-09-981-649A-11	Sequence 11, Appl
34	44	35.5	88	1	US-07-885-089B-13	Sequence 13, Appl
35	44	35.5	91	1	US-07-847-743B-15	Sequence 15, Appl
36	44	35.5	91	1	US-08-456-201-15	Sequence 15, Appl
37	44	35.5	91	1	US-08-330-161-13	Sequence 13, Appl
38	44	35.5	91	1	US-08-456-241-15	Sequence 13, Appl
39	44	35.5	91	1	US-08-440-401-13	Sequence 13, Appl
40	44	35.5	91	1	US-08-419-878B-13	Sequence 13, Appl
41	44	35.5	91	2	US-09-173-480-13	Sequence 13, Appl
42	44	35.5	91	2	US-10-022-609-13	Sequence 13, Appl
43	44	35.5	91	4	PCT-US92-04295A-15	Sequence 13, Appl
44	44	35.5	93	1	US-08-343-401A-4	Sequence 4, Appl
45	44	35.5	93	1	US-08-445-265A-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-198-452A-753
Sequence 753, Application US/09198452A
Patent No. 5332669
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment:
thereof and uses thereof, in particular for the diagnosis, prev.
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 753
LENGTH: 91
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-753

Query Match      100.0%; Score 124; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 66-13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TAFDPMKMLDGVCTYKGVQYXL 23
DB      15 TAFDPMKMLDGVCTYKGVQYXL 37

RESULT 2
US-09-438-185A-712
Sequence 712, Application US/09438185A
Patent No. 682071
GENERAL INFORMATION:
APPLICANT: Stephens, Richard
APPLICANT: Mitchell, Wayne
APPLICANT: Kaiman, Sue
APPLICANT: Davis, Ronald
TITLE OF INVENTION: The Regents of the University of California
Chlamydia Pneumoniae Genome Sequence
FILE REFERENCE: 018941-00041US
CURRENT APPLICATION NUMBER: US/09/438,185A
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/128,606
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 1074
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 712
LENGTH: 91
```

```

; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPn0710
US-09-438-185A-712

```

Query Match	100.0%;	Score 124;	DB 2;	Length 91;
Best Local Similarity	100.0%;	Pred. No. 6e-13;		
Matches 23; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

```
QY      1 TAFDENKMLDGVCTYVKGVQQYL 23
         |||||
Db     15 TAFDENKMLDGVCTYVKGVQQYL 37
```

RESULT 3
US-09-902-540-10545
; Sequence 10545, Application US/09902540

```

1 GENERAL INFORMATION:
2 APPLICANT: Goldman, Barry S.
3 APPLICANT: Hinkle, Gregory J.
4 APPLICANT: Slater, Steven C.
5 APPLICANT: Wiegand, Roger C.
6 TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
7 FILE REFERENCE: 38-10(15849)B
8 CURRENT APPLICATION NUMBER: US/09/902,540
9 CURRENT FILING DATE: 2001-07-10
10 PRIOR APPLICATION NUMBER: 60/217,883
11 PRIOR FILING DATE: 2000-07-10
12 NUMBER OF SEQ. ID NOS: 16825
13 SEQ ID NO 10545
14 LENGTH: 520

```

Query Match	40.3%	Score 50;	DB 2;	Length 520;
Best Local Similarity	56.2%	Pred. No. 7.3;		
Matches	9;	Conservative	1;	Mismatches 6;
				Indels 0;
				Gaps 0;

```
QY      1  TAPDFNKMLDGVCTYY  16
          |||| : ||||
Db      52  TTFDFGRSPDGTCTYY  67
```

RESULT 4
US-08-838-957A-23
; Sequence 23, Application US/08838957A

GENERAL INFORMATION:
APPLICANT: Breitman, Martin L.
APPLICANT: Rosant, Janet
APPLICANT: Dumont, Daniel J.
APPLICANT: Yamaguchi, Terry P.
TITLE OF INVENTION: No. 5998167e1 Receptor Tyrosine Kinases
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3T2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,957A
FILING DATE: 23-APR-1997
CLASSIFICATION: A55

ATTORNEY/AGENT INFORMATION:
NAME: Kurdydyk, Linda M.
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-2122
TELECOMMUNICATION INFORMATION:
TELEPHONE: (443) 364-7313

TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 23
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids

Query Match	38.7%	Score 48	DB 1	Length 44
Best Local Similarity	53.8%	Pred. No. 0.9		
Matches 7, Conservative	3	Mismatches	3	Gaps 0

Qy	10	DGVCTYVKGVQY	22
		: :	
Db	17	DGVCMYIKALDKY	29

```

RESULT 5
5332669-3
; Patent No. 5332669
; APPLICANT: DEWEIL, THOMAS F.
; TITLE OF INVENTION: PROSTATE-DERIVED MITOGEN
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/356,739
; FILING DATE: 24-MAR-1989
; SEQ ID NO.3:
; LENGTH: 53
5332669-3

```

Query Match	38.7%	Score	48	DB	6	Length	53
Best Local Similarity	50.0%	Pred. No.	1.1				
Matches	7	Conservative	4	Mismatches	3	Indels	0
						Gaps	0

```
QY      9 LDGVCITYVKGVQY 22
          ||||| |::: |
Db     16 LDGVCMIYIALDKY 29
```

RESULT 6
US-09-533-029-110
Sequence 110, Application US/09533029
Patent No. 666446
GENERAL INFORMATION:
APPLICANT: Heard, Jacqueline
APPLICANT: Broun, Pierre
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Keddie, James
APPLICANT: Pineda, Omaira
APPLICANT: Adam, Luc
APPLICANT: Samaha, Raymond
APPLICANT: Zhang, James
APPLICANT: Yu, Guo-Iiang
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Iliang, Cai-Zhong
APPLICANT: Reuber, Lynne
TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: NBI-010
CURRENT APPLICATION NUMBER: US/09/533,029
CURRENT FILING DATE: 2000-03-22
EARLIER APPLICATION NUMBER: 60/125,814
EARLIER FILING DATE: 1999-03-23

NUMBER OF SEQ ID NOS: 121
SOFTWARE: Patencin Ver. 2.1
SEQ ID NO 110
LENGTH: 371
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G1650
US-09-533-029-110

Query Match 36.3%; Score 45; DB 2; Length 371;
Best Local Similarity 43.5%; Pred. No. 33;
Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 TAFDFKMLDGVCTYKGVQY 23
DB 308 TKTDRSMLDVIYKVSLOSQI 330

RESULT 7
US-09-134-001C-3049
Sequence 3049, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3049
LENGTH: 427
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3049

Query Match 36.3%; Score 45; DB 2; Length 427;
Best Local Similarity 77.8%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 DFNKMLDGV 12
DB 286 DFNKMLDGI 294

RESULT 8
US-08-039-364-14
Sequence 14, Application US/08039364
Patent No. 581393
GENERAL INFORMATION:
APPLICANT: Klagsbrun, Michael
APPLICANT: Abraham, Judith A.
APPLICANT: Higashiyama, Shigeki
APPLICANT: Beener, Gail F.
TITLE OF INVENTION: HEPARIN BINDING MITOGEN WITH
TITLE OF INVENTION: HOMOLOG TO EPIDERMAL GROWTH FACTOR
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX

OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/039,364
FILING DATE: 15 JUN 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/598,082
FILING DATE: 16 OCT 1990
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 05162/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-039-364-14

Query Match 35.5%; Score 44; DB 1; Length 37;
Best Local Similarity 46.2%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 DGVCTYKGVQY 22
DB 12 DGVCTYKGVQY 24

RESULT 9
US-09-158-710-14
Sequence 14, Application US/09158710
Patent No. 6235884
GENERAL INFORMATION:
APPLICANT: Klagsbrun, Michael
APPLICANT: Abraham, Judith A.
APPLICANT: Higashiyama, Shigeki
APPLICANT: Beener, Gail F.
TITLE OF INVENTION: HEPARIN BINDING MITOGEN WITH HOMOLOG TO EPIDERMAL
TITLE OF INVENTION: GROWTH FACTOR (EGF)
FILE REFERENCE: 05162/002003
CURRENT APPLICATION NUMBER: US/09/158,710
PRIOR FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: US 08/039,364
EARLIER FILING DATE: 1993-06-15
EARLIER APPLICATION NUMBER: US 07/598,082
EARLIER FILING DATE: 1990-10-16
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 37
TYPE: PRT
ORGANISM: Homo sapiens
US-09-158-710-14

Query Match 35.5%; Score 44; DB 2; Length 37;
Best Local Similarity 46.2%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 DGVCTYKGVQY 22
DB 12 DGVCTYKGVQY 24

RESULT 10
US-08-168-091A-44
Sequence 44, Application US/08168091A
Patent No. 5663862

GENERAL INFORMATION:
APPLICANT: Falschbach, Gerald.
APPLICANT: Faller, Douglas R.
APPLICANT: Rosen, Kenneth M.
APPLICANT: Cortias, Gabriel
TITLE OF INVENTION: Neutrotrophic Factor
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE AND COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/168,091A
FILING DATE: 15-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/953,742
FILING DATE: 29-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: HMI-002CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-168-091A-44

Query Match 35.5%; Score 44; DB 1; Length 42;
Best Local Similarity 46.2%; Pred. No. 3.9;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 10 DGVCTYVXGVQY 22
Db 15 DGVCMYIEALDKY 27

RESULT 11
US-08-899-437-12
Sequence 12, Application US/08899437
Patent No. 6121415
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,437
FILING DATE: 24-Jul-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: hEGF.egf
LOCATION: 1-46
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-899-437-12

Query Match 35.5%; Score 44; DB 2; Length 46;
Best Local Similarity 46.2%; Pred. No. 4.3;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 10 DGVCTYVXGVQY 22
Db 16 DGVCMYIEALDKY 28

RESULT 12
US-09-126-121-12
Sequence 12, Application US/09126121
Patent No. 6252051
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: hEGF.egf
LOCATION: 1-46
IDENTIFICATION METHOD:

OTHER INFORMATION:
US-09-126-121-12

Query Match 35.5%; Score 44; DB 2; Length 46;
Best Local Similarity 46.2%; Pred. No. 4.3;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 DGVCYVKGVOQY 22
DB 16 DGVCMTIEALDKY 28

RESULT 13

US-08-753-007A-17
Sequence 17, Application US/08753007A

Patent No. 6074841

GENERAL INFORMATION:

APPLICANT: Gearling, David P.

APPLICANT: Busfield, Samantha J.

TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

TITLE OF INVENTION: AND USES THEREFOR

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSES: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/753,007A

CLASSIFICATION: 536

FILING DATE: 19-NOV-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/699,591

FILING DATE: 19-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Faase, J. Peter

REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 07334/022001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-542-5070

TELEFAX: 617-542-8906

TELEX:

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-753-007A-17

Query Match 35.5%; Score 44; DB 2; Length 47;

Best Local Similarity 46.2%; Pred. No. 4.4;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 DGVCYVKGVOQY 22

DB 17 DGVCMTIEALDKY 29

RESULT 14

US-09-398-496-17

Sequence 17, Application US/09398496

Patent No. 6133423

GENERAL INFORMATION:

APPLICANT: Gearling, David P.

APPLICANT: Busfield, Samantha J.

TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

TITLE OF INVENTION: AND USES THEREFOR

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSES: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/398,496

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/753,007

FILING DATE: 19-NOV-1996

APPLICATION NUMBER: 08/699,591

FILING DATE: 19-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Faase, J. Peter

REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 07334/022001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-542-5070

TELEFAX: 617-542-8906

TELEX:

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 47 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-398-496-17

Query Match 35.5%; Score 44; DB 2; Length 47;

Best Local Similarity 46.2%; Pred. No. 4.4;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 10 DGVCYVKGVOQY 22

DB 17 DGVCMTIEALDKY 29

RESULT 15

US-09-553-769-8

Sequence 8, Application US/09553769

Patent No. 6544759

GENERAL INFORMATION:

APPLICANT: Harari, Daniel

APPLICANT: Yarden, Yosef

TITLE OF INVENTION: NOVEL GROWTH FACTOR WHICH ACTS THROUGH E-DB-4 RECEPTOR TYROSINE

TITLE OF INVENTION: SEQUENCES ENCODING SAME AND USES THEREOF

FILE REFERENCE: 00/20522

CURRENT APPLICATION NUMBER: US/09/553,769

CURRENT FILING DATE: 2000-04-21

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn version 3.0

SEQ ID NO 8

LENGTH: 47

TYPE: PRT

ORGANISM: Homo sapiens

US-09-553-769-8

Query Match 35.5%; Score 44; DB 2; Length 47;

Best Local Similarity 46.2%; Pred. No. 4.4;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 10 DGVCTTYVKGVOQY 22

Db 17 DGVCMYIEALDKY 29

Search completed: January 23, 2006, 09:35:49
Job time : 17.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:32:06 ; Search time 60 Seconds
(without alignments)
160.168 Million cell updates/sec

Title: US-10-501-838A-10

Sequence: 1 TAFDFNKMIDGVCYVKGVOYL 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	100.0	23	4 US-10-665-184-10	Sequence 10, Appl
2	124	100.0	23	5 US-10-942-300-10	Sequence 10, Appl
3	124	100.0	23	5 US-10-501-838A-10	Sequence 10, Appl
4	124	100.0	91	4 US-10-289-762-753	Sequence 753, Appl
5	94	75.8	86	4 US-10-784-880-120	Sequence 120, Appl
6	50.5	40.7	45	4 US-10-424-589-214114	Sequence 214114, Appl
7	48.5	39.1	699	4 US-10-369-493-21251	Sequence 21251, Appl
8	47	37.9	74	4 US-10-424-599-252515	Sequence 252515, Appl
9	47	37.9	285	4 US-10-282-122A-52677	Sequence 52677, Appl
10	47	37.9	450	4 US-10-437-963-204288	Sequence 204288, Appl
11	47	37.9	475	4 US-10-437-963-168591	Sequence 168591, Appl
12	47	37.9	516	6 US-11-097-143-17850	Sequence 17850, Appl
13	47	37.9	712	4 US-10-425-115-301230	Sequence 301230, Appl
14	46	37.1	52	4 US-10-428-339-21	Sequence 21, Appl
15	46	37.1	68	4 US-10-106-698-5449	Sequence 5449, Appl
16	46	37.1	137	4 US-10-767-701-53692	Sequence 53692, Appl
17	46	37.1	415	4 US-10-369-493-20587	Sequence 20587, Appl
18	46	37.1	1548	4 US-10-437-963-179431	Sequence 179431, Appl
19	46	37.1	52	4 US-10-428-339-20	Sequence 20, Appl
20	45	36.3	145	4 US-10-767-701-35116	Sequence 35116, Appl
21	45	36.3	204	4 US-10-424-599-199200	Sequence 199200, Appl
22	45	36.3	371	3 US-09-533-029-110	Sequence 110, Appl
23	45	36.3	379	4 US-10-412-69898-932	Sequence 932, Appl
24	45	36.3	379	4 US-10-437-963-188768	Sequence 188768, Appl
25	45	36.3	427	4 US-10-724-972A-5883	Sequence 5883, Appl
26	45	36.3	431	4 US-10-425-114-53933	Sequence 53933, Appl
27	45	36.3	506	4 US-10-425-115-348627	Sequence 348627, Appl

28	45	36.3	659	4 US-10-425-115-250922	Sequence 250922, Appl
29	45	36.3	701	4 US-10-425-115-250924	Sequence 250924, Appl
30	45	36.3	861	4 US-10-437-963-16830	Sequence 16830, Appl
31	45	36.3	1139	4 US-10-369-493-6212	Sequence 6212, Appl
32	44	35.5	18	4 US-10-350-405-38	Sequence 38, Appl
33	44	35.5	46	3 US-09-817-647-12	Sequence 12, Appl
34	44	35.5	46	3 US-09-877-665-12	Sequence 12, Appl
35	44	35.5	46	4 US-10-136-572A-12	Sequence 12, Appl
36	44	35.5	46	4 US-10-215-862-12	Sequence 12, Appl
37	44	35.5	46	5 US-10-944-116-12	Sequence 12, Appl
38	44	35.5	46	5 US-11-035-787-12	Sequence 12, Appl
39	44	35.5	47	4 US-10-096-241-17	Sequence 8, Appl
40	44	35.5	47	4 US-10-240-441-8	Sequence 17, Appl
41	44	35.5	47	5 US-10-899-227-17	Sequence 13, Appl
42	44	35.5	48	4 US-10-201-945-13	Sequence 15, Appl
43	44	35.5	48	4 US-10-082-747A-15	Sequence 12, Appl
44	44	35.5	50	4 US-10-609-370-12	Sequence 2, Appl
45	44	35.5	50	6 US-11-018-188-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT 1
US-10-665-184-10
; Sequence 10, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-665-184-10

Query Match      100.0%; Score 124; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 5e-12; 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

Cy      1 TAFDFNKMIDGVCYVKGVOYL 23
Db      1 TAFDFNKMIDGVCYVKGVOYL 23

RESULT 2
US-10-942-300-10
; Sequence 10, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

```

PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-942-300-10

Query Match 100.0%; Score 124; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 5e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCITYKGVQOYL 23
Db 1 TAFDFNKMLDGVCITYKGVQOYL 23

RESULT 3

US-10-501-838A-10
; Sequence 10, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-501-838A-10

Query Match 100.0%; Score 124; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 5e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCITYKGVQOYL 23
Db 1 TAFDFNKMLDGVCITYKGVQOYL 23

RESULT 4

US-10-289-762-753
; Sequence 753, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 753
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-753

Query Match 100.0%; Score 124; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCITYKGVQOYL 23
Db 15 TAFDFNKMLDGVCITYKGVQOYL 37

RESULT 5

US-10-784-880-120
; Sequence 120, Application US/10784880
; Publication No. US20040208890A1
; GENERAL INFORMATION:
; APPLICANT: SUBTEL-SANDS, AGATHE
; APPLICANT: DAUTRY-VARSAT, ALICE
; TITLE OF INVENTION: SECRETED CHLAMYDIA POLYPEPTIDES, POLYNUCLEOTIDES CODING THEREOF
; FILE REFERENCE: 2491790S0
; CURRENT APPLICATION NUMBER: US/10/784,880
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US 60/448,879
; PRIOR FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 378
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 120
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Chlamydia psittaci
US-10-784-880-120

Query Match 75.8%; Score 94; DB 4; Length 86;
Best Local Similarity 73.9%; Pred. No. 9.3e-07;
Matches 17; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TAFDFNKMLDGVCITYKGVQOYL 23
Db 10 SAFNFMDLNGVCIXYGVQOYL 32

RESULT 6

US-10-424-599-214114
; Sequence 214114, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kowalic David K
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 214114
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_35371C.1.pep
US-10-424-599-214114

Query Match 40.7%; Score 50.5; DB 4; Length 45;
Best Local Similarity 64.7%; Pred. No. 2.7;
Matches 11; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 6 NKMLDGVCTYKGVQOYL 22
Db 7 NKNLQGLCS-VKGVSOY 22

RESULT 7

US-10-369-493-21251
; Sequence 21251, Application US/10369493
; Publication No. US2003023675A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xiandeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 21251
/ LENGTH: 699
/ TYPE: PRT
/ ORGANISM: Xenorhabdus nematophilus
/ US-10-369-493-21251

Query Match          39.1%; Score 48.5; DB 4; Length 699;
Best Local Similarity 47.8%; Pred. No. 1e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 1 TAFDFNKMLDGVCITYKGVQOYL 23
DB 268 TLFQNLGQLD---TFVRNVEQYL 267

RESULT 8
US-10-424-599-252515
/ Sequence 252515, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovalic, David K
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 252515
/ LENGTH: 74
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_70048C.1.pep
/ US-10-424-599-252515

Query Match          37.9%; Score 47; DB 4; Length 74;
Best Local Similarity 43.5%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 4; Indels 4; Gaps 2;

QY 3 FDFNKMLDGVCITYK--GVQOYL 23
DB 49 FDFGKKT-ICPYRIITGVQYV 69

RESULT 9
US-10-282-122A-52677
/ Sequence 52677, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zykied, Judith
/ APPLICANT: Wall, Daniel
```

```
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 52677
/ LENGTH: 285
/ TYPE: PRT
/ ORGANISM: Clostridium botulinum
/ US-10-282-122A-52677

Query Match          37.9%; Score 47; DB 4; Length 285;
Best Local Similarity 40.9%; Pred. No. 67;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 ADFPNKMLDGVCITYKGVQOYL 23
DB 254 ADDMNNALDKSIRYKQIEQYI 275

RESULT 10
US-10-437-963-204288
/ Sequence 204288, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 204288
/ LENGTH: 450
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_99390C.1.pep
/ US-10-437-963-204288
```

Query Match 37.9%; Score 47; DB 4; Length 450;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 5 FNIKMLDGVCTYKGVQ 20
|:|:|:|:|:|:
Db 143 FERRLADGVCDPVKLYE 158

RESULT 11
US-10-437-963-168591
; Sequence 168591, Application US/10437963
; Publication No. US20040123343A1

GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168591
; LENGTH: 475
; TYPE: PRT

ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_67091C.1.psp
US-10-437-963-168591

Query Match 37.9%; Score 47; DB 4; Length 475;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 DENKMLDGVCTY 15
|:|:|:|:|:|:
Db 398 DFDKMDGDCPLX 409

RESULT 12
US-11-097-143-17859
; Sequence 17859, Application US/11097143
; Publication No. US20050208558A1

GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.

TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04

PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19

PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12

PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60/191,637

; PRIOR FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 43008

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17859

; LENGTH: 516

; TYPE: PRT

; ORGANISM: DROSOPHILA

US-11-097-143-17859

Query Match 37.9%; Score 47; DB 6; Length 516;
Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FDFNKMLDGVCTY 15
|:|:|:|:|:|:
Db 345 FDFNRTSDRICTP 357

RESULT 13
US-10-425-115-301230
; Sequence 301230, Application US/10425115
; Publication No. US20040214272A1

GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 301230
; LENGTH: 712
; TYPE: PRT

ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(712)

OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_3798C.1.psp
US-10-425-115-301230

Query Match 37.9%; Score 47; DB 4; Length 712;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TAFDENKMLDGVCTYKGVQ 20
|:|:|:|:|:|:
Db 367 TVGEFRKRVDTLCELVRGMQ 386

RESULT 14
US-10-428-339-21
; Sequence 21, Application US/10428339
; Publication No. US20030228612A1

GENERAL INFORMATION:
; APPLICANT: KENWARD, Kimberly D.
; APPLICANT: SALEHUIZZAMAN, Shah

TITLE OF INVENTION: PRODUCTION OF RECOMBINANT EPIDERMAL
; TITLE OF INVENTION: GROWTH FACTOR IN PLANTS
; FILE REFERENCE: 07121.000502
; CURRENT APPLICATION NUMBER: US/10/428,339
; CURRENT FILING DATE: 2003-04-30

PRIOR APPLICATION NUMBER: 60/377,294
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 21
; LENGTH: 52

```

/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence; note =
/ OTHER INFORMATION: synthetic construct
US-10-428-339-21

```

```

Query Match          37.1%; Score 46; DB 4; Length 52;
Best Local Similarity 31.8%; Pred. No. 16;
Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      1 TAFDENKMLDGVCTYVKGVQY 22
      ::|||:|||||:|:|
Db      8 SSYDGYCLYNGVCMYIEAVDRY 29

```

```

RESULT 15
US-10-106-698-5449
/ Sequence 5449, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
/ FILE REFERENCE: PA005PI
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ PRIOR FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: Patent In Ver. 3.0
/ SEQ ID NO 5449
/ LENGTH: 68
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (62)_
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (63)_
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (66)_
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5449

```

```

Query Match          37.1%; Score 46; DB 4; Length 68;
Best Local Similarity 38.1%; Pred. No. 21;
Matches 8; Conservative 8; Mismatches 3; Indels 2; Gaps 1;

```

```

Qy      3 FDFRKMLDGVCTYVKGVQY 23
      |::|::|::|:|:|:|:|:|
Db      6 FHYNQIIEITC--VKGLQBYI 24

```

Search completed: January 23, 2006, 11:52:36
 Job time : 61 secs

This Page Blank (usps)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 23, 2006, 09:32:56 ; Search time 12.5546 Seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-11
Perfect score: 109
Sequence: 1 RALPLALAGLCQVAFAGDIS 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	51.4	305	JC5844	chitinase (EC 3.2.1.14) III Cl0501 - rice
2	54	49.5	236	T36399	outer membrane protein
3	49	45.0	35	I39369	conserved hypothetical
4	48	44.0	208	AH0199	ycfC protein - Bac
5	48	44.0	213	S19211	hypothetical protein
6	48	44.0	213	D90829	hypothetical protein
7	48	44.0	213	B85687	glutamate dehydrogenase
8	48	44.0	411	T49883	probable sigma-54
9	46.5	42.7	506	AC0845	probable sigma-54
10	46	42.2	409	A70647	probable sigma-54
11	45	41.3	215	AB0647	conserved hypothetical
12	45	41.3	455	C71345	probable amino acid
13	44	40.4	211	D84960	hypothetical protein
14	44	40.4	313	S75329	hypothetical protein
15	44	40.4	365	JU0319	hypothetical protein
16	44	40.4	386	AH0681	probable regulator
17	44	40.4	409	B95372	probable integrase
18	44	40.4	409	F95409	probable integrase
19	44	40.4	2157	A13009	peptide synthetase
20	44	40.4	2566	E98274	hypothetical protein
21	43.5	39.9	255	S31866	gamma-1 chain C
22	43	39.4	177	AB3269	outer membrane lip
23	43	39.4	181	B60738	insulin-like growth
24	43	39.4	202	T46523	probable cyclase-d
25	43	39.4	205	T64155	hypothetical protein
26	43	39.4	221	S58685	hypothetical protein
27	43	39.4	265	T08151	MDG homolog - Pyr
28	43	39.4	298	D75470	MD-repeat family P
29	43	39.4	318	D82742	dihydroxydipicolin

30	43	39.4	358	2	P70583	hypothetical prote
31	43	39.4	405	2	AD0276	probable ROK fami
32	43	39.4	406	2	AB5765	probable NAGC-like
33	43	39.4	406	2	D90916	probable NAGC-like
34	43	39.4	406	2	D64915	mLc protein - Bac
35	43	39.4	412	2	H75206	maltose transport
36	43	39.4	420	2	A54759	cytochrome ba(3) c
37	43	39.4	424	2	S46293	site-specific DNA-
38	43	39.4	481	2	D87489	NADH dehydrogenase
39	43	39.4	961	2	T32493	unc-45 protein - C
40	42.5	39.0	529	2	A65051	YgaA protein - Bac
41	42.5	39.0	529	2	E91074	probable 2-compone
42	42.5	39.0	537	2	D85919	probable 2-compone
43	42	38.5	137	2	G84174	hypothetical prote
44	42	38.5	190	2	S76635	hypothetical prote
45	42	38.5	207	2	H72056	Holliday junction

ALIGNMENTS

RESULT 1

JC5844
chitinase (EC 3.2.1.14) III Cl0501 - rice
N/Alternate names: 1,4-beta-poly-N-acetylglucosaminidase; chitodextrinase; poly-beta-gli
C/Species: Oryza sativa (rice)
C/Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C/Accession: JC5844
R/Nagasaki, H.; Yamamoto, K.; Shomura, A.; Koga-Ban, Y.; Takasuga, A.; Yano, M.; Minobe,
DNA Res. 4, 379-385, 1997
A/Title: Rice class III chitinase homologues isolated by random cloning of rice cDNAs.
A/Reference number: JC5844; MIMD:98162724; PMID:9501993
A/Accession: JC5844
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-305 <NAG>
A/Cross-references: UNIPROT:O49827; UNIPARC:UPI00000A8276; DDBJ:D55711; NID:92696226; P
C/Comment: This protein hydrolyzes poly-beta-1,4-N-acetyl glucosamine (chitin), and act
C/Superfamily: plant chitinase III
C/Keywords: glycosidase; hydrolase
F/30-305/Product: chitinase III Cl0501 1 #status predicted <MAT>
F/156/Active site: Glu #status predicted

Query Match 51.4%; Score 56; DB 2; Length 305;
Best Local Similarity 60.0%; Pred. No. 0.45;

Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 ILPLALAGLCQVAFAGDIS 22

DB 14 IMVVALAGLAGARAGDIA 33

RESULT 2

T36399
probable alcohol dehydrogenase (zinc-binding) - Streptomyces coelicolor (fragment)
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
R/Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Randleam, M.A.
submitted to the EMBL Data Library, August 1999
A/Reference number: Z21605
A/Accession: T36399
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-236 <NOR>
A/Cross-references: UNIPROT:O6CK64; UNIPARC:UPI0000174E49; EMBL:AL109987; PIDN:CAB53403
A/Experimental source: strain A3(2)
C/Genetics:
A/Genes: SCOEDB:SCF20.05
C/Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
Query Match 49.5%; Score 54; DB 2; Length 236;
Best Local Similarity 50.0%; Pred. No. 0.73;

Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 RAILPLALAGICQVARGD 20
: : : : :
Db 85 KVIYVHLITCGICRACRAGD 104

RESULT 3

139969
outer membrane protein A - Bacillus subtilis
C|Species: Bacillus subtilis
C|Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C|Accession: 139969
R|Ikemura, H.; Takagi, H.; Inouye, M.
J. Biol. Chem. 262, 7859-7864, 1987
A|Title: Requirement of pro-sequence for the production of active subtilisin E in Escher
A|Reference number: 139969; PMID:8722417; PMID:3108260
A|Accession: 139969
A|Status: translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-35 <RES>
A|Cross-references: UNIPROT:Q45665; UNIPARC:UPI00000B7287; GB:M16639; NID:9143521; PIDN:
A|Experimental source: strain M168, substrain PY79
A|Note: sequence was not translated in the genome sequence, reference number A69580
C|Genetics:
A|Gene: ompA
C|Superfamily: outer membrane protein A
C|Keywords: membrane protein

Query Match 45.0%; Score 49; DB 2; Length 35;
Best Local Similarity 59.1%; Pred. No. 0.76;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AILPLALAGICQVARGDIS 23
: : : : :
Db 5 AIAIVNALGAFATVAAQGISS 26

RESULT 4

AH0189
conserved hypothetical protein YPO1637 [imported] - Yersinia pestis (strain CO92)
C|Species: Yersinia pestis
C|Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C|Accession: AH0189
R|Patthill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
H., M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 523-527, 2001
A|Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A|Reference number: AB0001; PMID:21470413; PMID:11586360
A|Accession: AH0189
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-208 <KUR>
A|Cross-references: UNIPROT:Q8ZF06; UNIPARC:UPI00000CD828; GB:AL590842; PIDN:CAC90459.1;
C|Genetics:
A|Gene: YPO1637

Query Match 44.0%; Score 48; DB 2; Length 208;
Best Local Similarity 76.9%; Pred. No. 5.6;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IPLALAGICQVAR 17
: : : : :
Db 8 ITLALAGICQSR 20

RESULT 5

S19211
ycfC protease - Escherichia coli (strain K-12)
C|Species: Escherichia coli
C|Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C|Accession: S19211; A64858

R|Green, S.M.; Drabble, W.T.
submitted to the EMBL Data Library, May 1991
A|Description: Molecular analysis of the purB-phoP region of Escherichia coli K12.
A|Reference number: S19210
A|Accession: S19211
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-213 <GRE>

A|Cross-references: UNIPROT:P25746; UNIPARC:UPI000013A673; EMBL:X59307; NID:942582; PIDN
A|Experimental source: strain K-12
R|Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A|Title: The complete genome sequence of Escherichia coli K-12.
A|Reference number: A64720; PMID:97426617; PMID:9278503
A|Accession: A64858
A|Status: nucleic acid sequence not shown; translation not shown
A|Molecule type: DNA
A|Residues: 1-213 <BLAT>
A|Cross-references: UNIPARC:UPI000013A673; GB:A6000213; GB:U00096; NID:91787371; PIDN:AA
A|Experimental source: strain K-12, substrain MG1655
C|Genetics:
A|Gene: ycfC
A|Start codon: GTG

Query Match 44.0%; Score 48; DB 2; Length 213;
Best Local Similarity 76.9%; Pred. No. 5.7;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IPLALAGICQVAR 17
: : : : :
Db 8 ITLALAGICQSR 20

RESULT 6

D90829
hypothetical protein Ecol604 [imported] - Escherichia coli (strain O157:H7, substrain RI
C|Species: Escherichia coli
C|Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C|Accession: D90829
R|Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasaawara, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hatlori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A|Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A|Reference number: A99629; PMID:21156231; PMID:11258796
A|Accession: D90829
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-213 <HAV>
A|Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:BA000007; PIDN:BA835027.1;
A|Experimental source: strain O157:H7, substrain R1MD 0509952
C|Genetics:
A|Gene: Ecol604

Query Match 44.0%; Score 48; DB 2; Length 213;
Best Local Similarity 76.9%; Pred. No. 5.7;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IPLALAGICQVAR 17
: : : : :
Db 8 ITLALAGICQSR 20

RESULT 7

B85687
hypothetical protein ycfC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C|Species: Escherichia coli
C|Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C|Accession: B85687
R|Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hiller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A|Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: AB5480; MUID:21074935; PMID:11206551
 A:Accession: B85687
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-213 <STO>
 A:Cross-references: UNIPROT:Q8X736; UNIPARC:UPI0000000AD2; GB:AE005174; NID:g12514786; F
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: ycfC

Query Match 44.0%; Score 48; DB 2; Length 213;
 Best Local Similarity 76.9%; Pred. No. 5.7;
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IPLALAGLCQVAR 17
 |||||:|||||
 DB 8 IITLALGICQVSAR 20

RESULT 8
 T49883
 glutamate dehydrogenase 2 - Arabidopsis thaliana
 N:Alternate names: protein T211.150
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 05-Oct-2004
 C:Accession: T49883
 R:Byers, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
 Submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24493
 A:Accession: T49883
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-411 <BEV>
 A:Cross-references: UNIPROT:Q39946; UNIPARC:UPI00001292DA; EMBL:AL163912; GSPDB:GN00063;
 A:Experimental source: cultivar Columbia; BAC clone T211
 C:Genetics:
 A:Gene: ATSP:T211.150
 A:Map position: 5
 A:Introns: 36/3; 74/3; 173/3; 212/3; 296/2; 321/3
 C:Superfamily: glutamate dehydrogenase

Query Match 44.0%; Score 48; DB 2; Length 411;
 Best Local Similarity 42.9%; Pred. No. 10;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 IITLALAGLCQVARADISS 23
 |||||:|||||
 DB 284 VLIPLCALGVLNKNMGDVRA 304

RESULT 9
 AC0845
 probable sigma-54-dependent transcriptional regulator STY2961 [imported] - Salmonella ent
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A:Note: this species has also been called Salmonella typh
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 05-Oct-2004
 C:Accession: AC0845
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
 , S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AC0845
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-506 <PAR>
 A:Cross-references: UNIPARC:UPI000005A33F; GB:AL513382; PIDN:CAD05946.1; PID:g16503917;
 C:Genetics:
 A:Gene: STY2961
 C:Superfamily: response regulator (sigma54-dependent transcriptional activator), FliA ty

Query Match 42.7%; Score 46.5; DB 2; Length 506;
 Best Local Similarity 35.3%; Pred. No. 21;
 Matches 12; Conservative 3; Mismatches 2; Indels 17; Gaps 1;

QY 5 IPLALAGLCQ-----VARAGDI 21
 |||||:|||||
 DB 52 IITLALAGLCQVAGRRFTLEGHPRLEALRAGDV 85

RESULT 10
 A70647
 probable PPR protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: A70647
 R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
 i Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
 A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: A70647
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-409 <COL>
 A:Cross-references: UNIPROT:P95182; UNIPARC:UPI00000CCBBF; GB:Z83667; GB:AL123456; NID:
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: PPR

Query Match 42.2%; Score 46; DB 2; Length 409;
 Best Local Similarity 61.1%; Pred. No. 21;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 IPLALAGLCQVARADISS 23
 |||||:|||||
 DB 279 IITLALAGLCQVAGRTNSA 296

RESULT 11
 AB0647
 conserved hypothetical protein STY1273 [imported] - Salmonella enterica subsp. enterica
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A:Note: this species has also been called Salmonella typh
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB0647
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
 , S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AB0647
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-215 <PAR>
 A:Cross-references: UNIPARC:UPI0000059F77; GB:AL513382; PIDN:CAD08357.1; PID:g16502402;
 C:Genetics:
 A:Gene: STY1273

Query Match 41.3%; Score 45; DB 2; Length 215;
 Best Local Similarity 69.2%; Pred. No. 17;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 IPLALAGLCQVAR 17
 |||||:|||||
 DB 10 IITLALGICQVSAR 22

RESULT 12
 C71345

probable amino acid ABC transporter, permease protein (brac) - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
 C:Accession: C71345
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwilt
 Iron, J.; Khalil, H.; Richardson, D.; Howell, J.R.; Chidambaram, M.; Uetzerback, T.; McD
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A11250; PMID:9832770; PMID:965876
 A:Accession: C71345
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-455 <CON>
 A:Cross-references: UNIPROT:O83289; UNIPARC:UPI00000D3223; GB:AB001208; GB:AB000520; NID
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0365
 C:Superfamily: branched-chain amino acid transport system II carrier protein braz

Query Match 41.3%; Score 45; DB 2; Length 455;
 Best Local Similarity 55.0%; Pred. No. 33;
 Matches 11; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

OY 2 AILPLALAGLCQVARAGDI 21
 |||||:|||||:
 Db 48 AVLIP-INGLITTHAGGI 65

RESULT 13
 D84960
 hypothetical protein [imported] - Buchnera sp. (strain APS)
 C:Species: Buchnera sp.
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: D84960
 R:Shigenobu, S.; Watanabe, H.; Hatori, M.; Sakaki, Y.; Ishikawa, H.
 Nature 407, 81-86, 2000
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
 A:Reference number: A84930; PMID:20445173; PMID:10993077
 A:Accession: D84960
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-211 <STO>
 A:Cross-references: UNIPARC:UPI000005E84F; GB:AP000398; GSPDB:GN00144
 A:Experimental source: strain APS
 C:Genetics:
 A:Gene: ycfC; BU262

Query Match 40.4%; Score 44; DB 2; Length 211;
 Best Local Similarity 69.2%; Pred. No. 24;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 LILPLALAGLCQVA 16
 |||||:|||||:
 Db 6 LITLSIAGICQSA 18

RESULT 14
 S75329
 hypothetical protein sll1571 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S75329
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 S.
 A:Reference number: S74322; PMID:97061201; PMID:8905231
 A:Accession: S75329
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A:Residues: 1-113 <KAN>
 A:Cross-references: UNIPROT:P73216; UNIPARC:UPI00000C0C0D1; EMBL:D90904; GB:AB001339; NID
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: Synechocystis hypothetical protein sll1571

Query Match 40.4%; Score 44; DB 2; Length 113;
 Best Local Similarity 52.9%; Pred. No. 34;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 LILPLALAGLCQVARAG 19
 |||||:|||||:
 Db 44 LILPFAITGLVWLAIAG 60

RESULT 15
 JU0319
 hypothetical 43.5K protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
 C:Accession: JU0319
 R:Nottingham, C.; Gallatin, W.M.; St. John, T.
 Gene 95, 279-284, 1990
 A:Title: Lymphocyte HEV adhesion variants differ in the expression of multiple gene sequ
 A:Reference number: JU0319; PMID:91065542; PMID:2174398
 A:Accession: JU0319
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-365 <NOT>
 A:Cross-references: UNIPROT:P42209; UNIPARC:UPI00000279D0; GB:M37030; NID:G193839; PIDN:
 C:Superfamily: Saccharomyces cerevisiae cell division control protein CDC10

Query Match 40.4%; Score 44; DB 2; Length 365;
 Best Local Similarity 53.3%; Pred. No. 39;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 IPLAIAGLCQVARAG 19
 |||||:|||||:
 Db 222 IPFAVGSCEVVRDG 236

Search completed: January 23, 2006, 10:43:05
 Job time : 13.5546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:35:50 / Search time 21.0917 Seconds
(without alignments)
90.156 Million cell updates/sec

Title: US-10-501-838a-11
Perfect score: 109
Sequence: 1 RALPLALAGLCQVARADISS 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5.COMB.pep.*
2: /cgn2_6/prodata/1/1aa/6.COMB.pep.*
3: /cgn2_6/prodata/1/1aa/H.COMB.pep.*
4: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep.*
5: /cgn2_6/prodata/1/1aa/RE.COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfill1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	47.7	305	2	US-09-178-610-6
2	49	45.0	217	2	US-09-489-039A-7789
3	48	44.0	239	2	US-09-270-767-42453
4	47	43.1	131	2	US-09-553-498-4
5	47	43.1	131	2	US-09-618-869-4
6	47	43.1	216	2	US-09-543-681A-7993
7	47	43.1	232	2	US-09-553-498-6
8	47	43.1	232	2	US-09-618-869-6
9	47	43.1	339	2	US-09-553-498-2
10	47	43.1	339	2	US-09-618-869-2
11	46	42.2	233	2	US-09-471-376-1333
12	46	42.2	233	2	US-08-812-586-45
13	46	42.2	233	2	US-09-535-832A-42
14	45.5	41.7	532	2	US-09-489-039A-10011
15	45	41.3	23	1	US-08-215-138-1
16	45	41.3	23	1	US-08-407-344-1
17	44	40.4	44	2	US-08-905-223-440
18	44	40.4	115	2	US-09-553-769-2
19	44	40.4	115	2	US-09-640-041-2
20	44	40.4	211	2	US-09-774-639-153
21	44	40.4	367	2	US-09-032-372-4
22	44	40.4	711	1	US-08-235-838-7
23	44	40.4	711	1	US-08-465-473B-7
24	43.5	39.9	834	1	US-08-677-734A-9
25	43.5	39.9	834	1	US-08-677-734A-10
26	43.5	39.9	834	2	US-09-097-053-9
27	43.5	39.9	834	2	US-09-097-053-10

28	43	39.4	343	2	US-09-902-540-13926	Sequence 13926, A
29	43	39.4	406	1	US-08-155-906-2	Sequence 2, Appl1
30	43	39.4	429	2	US-09-489-039A-13738	Sequence 13738, A
31	42	38.5	180	2	US-09-645-415A-8	Sequence 8, Appl1
32	42	38.5	207	2	US-09-198-452A-658	Sequence 658, App
33	42	38.5	209	2	US-09-438-185A-622	Sequence 622, App
34	42	38.5	303	2	US-09-902-540-14321	Sequence 14321, A
35	42	38.5	455	2	US-09-270-767-45790	Sequence 45790, A
36	42	38.5	761	2	US-09-198-452A-335	Sequence 335, App
37	42	38.5	902	2	US-09-438-185A-319	Sequence 319, App
38	41.5	38.1	259	2	US-09-540-236-3064	Sequence 3064, App
39	41	37.6	21	1	US-08-449-287-20	Sequence 20, Appl1
40	41	37.6	21	1	US-09-003-081-8	Sequence 8, Appl1
41	41	37.6	21	2	US-08-648-262-8	Sequence 8, Appl1
42	41	37.6	21	2	US-08-648-263-8	Sequence 8, Appl1
43	41	37.6	21	2	US-08-840-713-48	Sequence 48, Appl1
44	41	37.6	21	2	US-09-904-196B-3	Sequence 3, Appl1
45	41	37.6	21	2	US-09-230-233A-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-178-610-6
Sequence 6, Application US/09178610
Patent No. 6124126
GENERAL INFORMATION:
APPLICANT: ITO, YOSHIFUMI
APPLICANT: PARK, SEUNG-MOON
APPLICANT: NAM HAI, TRUONG
TITLE OF INVENTION: COMPLEMENTARY DNA FOR RICE CHITINASE HAVING LYTC
TITLE OF INVENTION: ACTIVITY AGAINST MOULDS AND BACTERIA, AND VECTOR
TITLE OF INVENTION: CONTAINING SAID COMPLEMENTARY DNA AND TRANSFORMANT
FILE REFERENCE: 8361-0006-0X
CURRENT APPLICATION NUMBER: US/09/178,610
CURRENT FILING DATE: 1998-10-26
EARLIER APPLICATION NUMBER: JP 123905
EARLIER FILING DATE: 1998-04-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 305
TYPE: PRT
ORGANISM: rice plant (Oryza sativa L. Nipponbare)
US-09-178-610-6

Query Match      47.7%; Score 52; DB 2; Length 305;
Best Local Similarity 55.0%; Pred. No. 4.1; Mismatches 0; Gaps 0;
Matches 11; Conservative 4; Indels 0;

Oy      3 ILPLALAGLCQVARADIS 22
Db      14 IMYVALAGLAAGTRAGDIA 33

RESULT 2
US-09-489-039A-7789
Sequence 7789, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7789
LENGTH: 217
TYPE: PRT
```



```

/ GENERAL INFORMATION:
/ APPLICANT: Ambrosius, Dorothee
/ APPLICANT: Rudolph, Rainer
/ APPLICANT: Schaeffner, Joerg
/ APPLICANT: Schwarz, Elisabeth
/ TITLE OF INVENTION: Process for the production of naturally folded and secreted proteins
/ FILE REFERENCE: Case 20379
/ CURRENT APPLICATION NUMBER: US/09/553,498
/ CURRENT FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: EP99107412.1
/ PRIOR FILING DATE: 1999-04-26
/ NUMBER OF SEQ ID NOS: 10
/ SEQ ID NO 2
/ LENGTH: 399
/ TYPE: PRT

```

```

1 FILE REFERENCE: GENSET.0250CPI
2
3 CURRENT APPLICATION NUMBER: 09/09/471, 276
4
5 CURRENT FILING DATE: 1999-12-21
6
7 EARLIER APPLICATION NUMBER: 09/057, 719
8
9 EARLIER FILING DATE: 1998-04-09
10
11 EARLIER APPLICATION NUMBER: 09/069, 047
12
13 EARLIER FILING DATE: 1998-04-28
14
15 EARLIER APPLICATION NUMBER: PCT/1999/00714
16
17 EARLIER FILING DATE: 1999-04-09
18
19 NUMBER OF SEQ ID NOS: 1622
20
21 SOFTWARE: Patent.pm
22
23 SEQ ID NO 1333
24
25 LENGTH: 39
26
27 TYPE: prt
28
29 ORGANISM: Homo sapiens
30
31 FEATURES:
32
33 NAME/KEY: SIGNAL

```

LOCATION: -30.-1
US-09-471-276-1333

Query Match 42.2%; Score 46; DB 2; Length 39;
Best Local Similarity 53.3%; Pred. No. 3.7;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 ILPLALAGLCQVAR 17
DB 1 MLPLAMAGRCYAK 15

RESULT 12
US-08-812-586-45
Sequence 45, Application US/08812586
Patent No. 6048704
GENERAL INFORMATION:
APPLICANT: Martin David Tilson
TITLE OF INVENTION: PURIFIED AND RECOMBINANT ANTIGENIC
TITLE OF INVENTION: PROTEINS ASSOCIATED WITH ABDOMINAL AORTIC ANEURYSM (AAA)
TITLE OF INVENTION: DISEASE, AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812.586
FILING DATE: 07-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/53862-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-812-586-45

Query Match 42.2%; Score 46; DB 2; Length 233;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 AILPLALAGLCQVARAGDIS 22
DB 5 AIAIYVALAGFATVAAQAAELT 25

RESULT 13
US-09-535-832A-42
Sequence 42, Application US/09535832A
GENERAL INFORMATION:
APPLICANT: Tilson, Martin David
TITLE OF INVENTION: Purified and Recombinant Antigenic Proteins Associated
TITLE OF INVENTION: with Abdominal Aortic Aneurysm (AAA) Disease, and
TITLE OF INVENTION: Diagnostic and Therapeutic use thereof
FILE REFERENCE: 53862-AZ

CURRENT APPLICATION NUMBER: US/09/535,832A
CURRENT FILING DATE: 2000-03-28
NUMBER OF SEQ ID NOS: 57
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 42
LENGTH: 233
TYPE: PRT
ORGANISM: Homo sapiens
US-09-535-832A-42

Query Match 42.2%; Score 46; DB 2; Length 233;
Best Local Similarity 47.6%; Pred. No. 25;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 AILPLALAGLCQVARAGDIS 22
DB 5 AIAIYVALAGFATVAAQAAELT 25

RESULT 14
US-09-489-039A-10011
Sequence 10011, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10011
LENGTH: 532
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10011

Query Match 41.7%; Score 45.5; DB 2; Length 532;
Best Local Similarity 32.4%; Pred. No. 72;
Matches 11; Conservative 4; Mismatches 2; Indels 17; Gaps 1;

QY 5 IPLALAGCQ-----VARAGDI 21
DB 68 VPLAIDGLADVLGRRPALGPRLEAIARAGDV 101

RESULT 15
US-08-215-138-1
Sequence 1, Application US/08215138
Patent No. 5470719
GENERAL INFORMATION:
APPLICANT: Meng, Shi-Yuan
APPLICANT: Morris, Charles F.
APPLICANT: Tsai, Larry B.
TITLE OF INVENTION: Enhanced Secretion of Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Amgen Inc., U.S. Patent Operations/NAO
STREET: 1840 Dehavenland Drive
CITY: Thousand Oaks
STATE: CA
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/215,138
FILING DATE:

CLASSIFICATION: 530
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 23 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-215-138-1

Query Match 41.3%; Score 45; DB 1; Length 23;
 Best Local Similarity 61.1%; Pred. No. 3;
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 RAIIPLALAGLCQVABA 18
 DB 6 RAIIAIVALLAGPATVANA 23

Search completed: January 23, 2006, 10:46:47
 Job time : 22.0917 secs

This Page Blank (Copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comugen Ltd.

OM protein - protein search, using bw model

Run on: January 23, 2006, 11:53:39 ; Search time 3.21397 Seconds
(without alignments)
72.521 Million cell updates/sec

Title: US-10-501-838a-11

Sequence: 1 RAILPLPALAGLCQVARGADISS 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New.*
1: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubppaa/PC7_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubppaa/US05_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubppaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	40.4	210	6	US-10-986-501-115
2	42	38.5	180	7	US-11-082-544-8
3	41	37.6	266	7	US-11-082-544-10
4	41	37.6	445	6	US-10-063-703-32
5	41	37.6	445	7	US-11-102-240-32
6	40	36.7	427	6	US-10-995-561-816
7	40	36.7	427	6	US-10-995-561-818
8	39	35.8	123	6	US-10-432-483-4
9	39	35.8	829	7	US-11-194-246-296
10	39	35.8	6893	7	US-11-205-109-14
11	39	35.8	8695	7	US-11-205-109-15
12	38	34.9	357	7	US-11-108-528-60
13	38	34.9	407	7	US-11-092-140-4
14	38	34.9	459	7	US-11-143-980-65
15	37.5	34.4	512	6	US-10-467-657-6780
16	37	33.9	92	7	US-11-116-144-170
17	37	33.9	269	7	US-11-074-176-210
18	37	33.9	291	6	US-10-467-657-5640
19	37	33.9	360	6	US-10-858-730-110
20	37	33.9	399	6	US-10-821-234-1031
21	37	33.9	468	6	US-10-467-657-1112
22	36.5	33.5	104	7	US-11-123-896-261
23	36.5	33.5	179	6	US-10-467-657-2566
24	36.5	33.5	307	6	US-10-467-657-2792
25	36.5	33.5			Sequence 2792, Ap

25	36.5	33.5	307	6	US-10-467-657-6738	Sequence 6738, Ap
27	36.5	33.5	480	6	US-10-821-234-1465	Sequence 1465, Ap
28	36	33.0	131	7	US-11-084-591-3	Sequence 3, Appl
29	36	33.0	142	7	US-11-186-284-105	Sequence 105, Appl
30	36	33.0	204	7	US-11-055-822-468	Sequence 468, App
31	36	33.0	260	7	US-11-055-822-514	Sequence 514, App
32	36	33.0	331	7	US-11-129-143-87	Sequence 87, Appl
33	36	33.0	345	6	US-10-467-657-7598	Sequence 7598, Ap
34	36	33.0	367	6	US-10-821-234-1569	Sequence 1569, Ap
35	36	33.0	473	7	US-11-075-185-12	Sequence 12, Appl
36	36	33.0	483	7	US-11-113-424-24	Sequence 24, Appl
37	36	33.0	491	7	US-11-205-109-21	Sequence 21, Appl
38	36	33.0	524	7	US-11-113-424-64	Sequence 64, Appl
39	36	33.0	804	6	US-10-467-657-5508	Sequence 57, Appl
40	36	33.0	943	6	US-10-467-657-5508	Sequence 5508, Ap
41	36	33.0	1302	7	US-11-090-439-42	Sequence 42, Appl
42	35.5	32.6	1892	7	US-11-075-185-6	Sequence 6, Appl
43	35	32.1	248	7	US-11-080-628-23	Sequence 23, Appl
44	35	32.1	279	7	US-11-055-822-614	Sequence 614, App
45	35	32.1	312	6	US-10-858-730-117	Sequence 117, App

ALIGNMENTS

RESULT 1
US-10-986-501-115
; Sequence 115, Application US/10986501
; Publication No. US20050244845A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCES: P2013P201
; CURRENT APPLICATION NUMBER: US/10/986,501
; PRIOR FILING DATE: 2004-11-12
; PRIOR APPLICATION NUMBER: US/10/621,363
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/969,730
; PRIOR FILING DATE: 2001-10-06
; PRIOR APPLICATION NUMBER: 09/774,639
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/238,291
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,732
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,366
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,364
; PRIOR FILING DATE: 1997-08-19
; Remaining PCT ID Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 115
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-986-501-115
Query Match 40.4%; Score 44; DB 6; Length 210;
Best Local Similarity 47.4%; Pred. No. 3.5;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
OY 1 RAILPLPALAGLCQVARG 19

Db 2 RCTTPTMLRALAQAARAG 20

RESULT 2
US-11-082-544-8

Sequence 8, Application US/11082544
Publication No. US20050249706A1
GENERAL INFORMATION:
APPLICANT: Bermudes, G.
APPLICANT: King, I.
APPLICANT: Clairmont, C.
APPLICANT: Lin, S.
APPLICANT: Belcourt, M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: TUMOR-TARGETED DELIVERY OF EFFECTOR MOLECULES
FILE REFERENCE: 8002-059
CURRENT APPLICATION NUMBER: US/11/082,544
CURRENT FILING DATE: 2005-03-17
PRIOR APPLICATION NUMBER: US/09/645,415
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/157,581
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: 60/157,637
PRIOR FILING DATE: 1999-10-04
NUMBER OF SEQ ID NOS: 61
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 180
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion construct
US-11-082-544-8

Query Match 38.5%; Score 42; DB 7; Length 180;
Best Local Similarity 50.0%; Pred. No. 6.2;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 2 AIIPLALAGLCQVAPAGDIS 23

Db 5 AIAIIVALAGPATVVAQAHMVR 26

RESULT 3
US-11-082-544-10

Sequence 10, Application US/11082544
Publication No. US20050249706A1
GENERAL INFORMATION:
APPLICANT: Bermudes, G.
APPLICANT: King, I.
APPLICANT: Clairmont, C.
APPLICANT: Lin, S.
APPLICANT: Belcourt, M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: TUMOR-TARGETED DELIVERY OF EFFECTOR MOLECULES
FILE REFERENCE: 8002-059
CURRENT APPLICATION NUMBER: US/11/082,544
CURRENT FILING DATE: 2005-03-17
PRIOR APPLICATION NUMBER: US/09/645,415
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/157,581
PRIOR FILING DATE: 1999-10-04
PRIOR APPLICATION NUMBER: 60/157,637
PRIOR FILING DATE: 1999-10-04
NUMBER OF SEQ ID NOS: 61
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 266
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion construct

US-11-082-544-10

Query Match 37.6%; Score 41; DB 7; Length 266;
Best Local Similarity 58.8%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 AIIPLALAGLCQVAPARA 18
Db 5 AIAIIVALAGPATVVAQA 21

RESULT 4
US-10-063-703-32

Sequence 32, Application US/10063703
Publication No. US20060008901A1
GENERAL INFORMATION:
APPLICANT: Baton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Macanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACTS ENCODING THE SAME
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,703
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 32
LENGTH: 445
TYPE: PRT
ORGANISM: Homo Sapien
US-10-063-703-32

Query Match 37.6%; Score 41; DB 6; Length 445;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 2 AIIPLALAGLCQVAPAGDI 21

Db 141 AVILSLGTRGAGAVMEGGQV 160

RESULT 5
US-11-102-240-32

Sequence 32, Application US/1102240
Publication No. US20050260647A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
TITLE OF INVENTION: ESOPHAGEAL TUMOR
FILE REFERENCE: P3230R1C106C
CURRENT APPLICATION NUMBER: US/11/102,240
CURRENT FILING DATE: 2005-04-08
PRIOR APPLICATION NUMBER: 10/063662
PRIOR FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 10/06867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/1170262
PRIOR FILING DATE: 199-12-09
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 32
LENGTH: 445

;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-11-102-240-32

Query Match 37.6%; Score 41; DB 7; Length 445;
Best Local Similarity 40.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 ALPLALAGLCQVAPAGDI 21
DB 141 AVLSTGTAGAAWMEGGGV 160

RESULT 6
US-10-995-561-816
; Sequence 816, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001359
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 816
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-816

Query Match 36.7%; Score 40; DB 6; Length 427;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 4 LIPLALAGLCQVAPAGDISS 23
DB 367 VVPTCLRALVEMARVGASS 386

RESULT 7
US-10-995-561-818
; Sequence 818, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001359
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 818
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-818

Query Match 36.7%; Score 40; DB 6; Length 427;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 4 LIPLALAGLCQVAPAGDISS 23
DB 367 VVPTCLRALVEMARVGASS 386

RESULT 8
US-10-432-483-4

; Sequence 4, Application US/10432483
; Publication No. US20050260699A1
; GENERAL INFORMATION:
; APPLICANT: desouza, Mervyn L.
; APPLICANT: Jensen, Holly
; APPLICANT: Schroeder, William A.
; APPLICANT: Gokarn, Ravi R.
; TITLE OF INVENTION: CAROTENOID BIOSYNTHESIS
; FILE REFERENCE: 12799-0020U1
; CURRENT APPLICATION NUMBER: US/10/432,483
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: PCT/US01/43906
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/252,749
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Agromyces mediolanus
US-10-432-483-4

Query Match 35.8%; Score 39; DB 6; Length 123;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 PLALAGLCQVAPAG 19
DB 80 PVFLAFLCQLAMVG 93

RESULT 9
US-11-194-246-296
; Sequence 296, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Motz, John
; APPLICANT: Treped, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND ME
; TITLE OF INVENTION: USE
; FILE REFERENCE: 00592.US1 (MAR 268 05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274,586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 296
; LENGTH: 829
; TYPE: PRT
; ORGANISM: HAEMOPHILUS INFLUENZAE
US-11-194-246-296

Query Match 35.8%; Score 39; DB 7; Length 829;
Best Local Similarity 56.2%; Pred. No. 90;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 IPLALAGLCQVAPAGD 20
DB 564 IPVFLGLSGVPAGD 579

RESULT 10
US-11-205-109-14
; Sequence 14, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel

```

; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
; FILE REFERENCE: 3002-205
; CURRENT APPLICATION NUMBER: US/11/205,109
; CURRENT FILING DATE: 2005-08-17
; PRIOR APPLICATION NUMBER: US/09/976,059
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/239,924
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; LENGTH: 6893
; TYPE: PRT
; ORGANISM: Actinoplanes sp.
US-11-205-109-14

Query Match      35.8%; Score 39; DB 7; Length 6893;
Best Local Similarity 42.9%; Pred. No. 8e+02;
Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy      3 ILIPLAAGLCQVAPAGDISS 23
Db      745 VLDPSCFAGLTREVTGDAVS 765

RESULT 11
US-11-205-109-15
; Sequence 15, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
; FILE REFERENCE: 3002-205
; CURRENT APPLICATION NUMBER: US/11/205,109
; CURRENT FILING DATE: 2005-08-17
; PRIOR APPLICATION NUMBER: US/09/976,059
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/239,924
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 8695
; TYPE: PRT
; ORGANISM: Actinoplanes sp.
US-11-205-109-15

Query Match      35.8%; Score 39; DB 7; Length 8695;
Best Local Similarity 42.9%; Pred. No. 1e+03;
Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy      3 ILIPLAAGLCQVAPAGDISS 23
Db      3889 VLDPSCFAGLTREVTGDAVS 3909

RESULT 12
US-11-108-528-60
; Sequence 60, Application US/11108528
; Publication No. US20050261189A1
; GENERAL INFORMATION:
; APPLICANT: Larsen, Glenn
; APPLICANT: Marvin, Martha
; APPLICANT: Li, Dean Y.
; APPLICANT: Wang, Elizabeth
; APPLICANT: Chen, C. M. Amy
; APPLICANT: Shamah, Steven M.
; TITLE OF INVENTION: METHODS OF PROMOTING CARDIAC CELL
; FILE REFERENCE: HYDR-P01-041
```

```

; CURRENT APPLICATION NUMBER: US/11/108,528
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: US 60/563,137
; PRIOR FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/598,368
; PRIOR FILING DATE: 2004-08-02
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-108-528-60

Query Match      34.9%; Score 38; DB 7; Length 357;
Best Local Similarity 80.0%; Pred. No. 54;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      7 LALAGLCOVA 16
Db      7 LALAGLCIVA 16

RESULT 13
US-11-092-140-4
; Sequence 4, Application US/11092140
; Publication No. US20050262590A1
; GENERAL INFORMATION:
; APPLICANT: Subramaniam, S.; Slater, S.; Karberg, K.; Chen, R.; Valentin, H.; Wong, Y.
; TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synthes
; FILE REFERENCE: 16515.054
; CURRENT APPLICATION NUMBER: US/11/092,140
; CURRENT FILING DATE: 2005-03-29
; PRIOR APPLICATION NUMBER: US/09/688,069
; PRIOR FILING DATE: 2000-10-14
; NUMBER OF SEQ ID NOS: 114
; SEQ ID NO 4
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-11-092-140-4

Query Match      34.9%; Score 38; DB 7; Length 407;
Best Local Similarity 50.0%; Pred. No. 62;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy      2 AILIPLAAGLC 13
Db      284 SIVPLYLISGVC 295

RESULT 14
US-11-143-980-65
; Sequence 65, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: He, Min
; APPLICANT: Hucul, John
; APPLICANT: Haili, Bradley A.
; APPLICANT: Wagenaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Sumners, Mia
; APPLICANT: Kulowski, Kerry
; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; PRIOR FILING DATE: 2004-06-03
```

/ NUMBER OF SEQ ID NOS: 72
 / SOFTWARE: PatentIn version 3.3
 / SEQ ID NO: 65
 / LENGTH: 459
 / TYPE: PRT
 / ORGANISM: Streptomyces sp.
 US-11-143-980-65

Query Match 34.9%; Score 38; DB 7; Length 459;
 Best Local Similarity 47.4%; Pred. No. 70;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 RAILIPLALAGLQVVARAG 19
 ||:|||||:||||
 Db 370 RATEVPLAEGRLELQRAQ 388

RESULT 15
 US-10-467-657-6780
 / Sequence 6780, Application US/10467657
 / Publication No. US20050260581A1
 / GENERAL INFORMATION:
 / APPLICANT: CHIRON SPA
 / APPLICANT: FONTANA Maria Rita
 / APPLICANT: PIZZA Mariagrazia
 / APPLICANT: MASIGNANI Vega
 / APPLICANT: MONACI Elisabetta
 / TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 / FILE REFERENCE:
 / CURRENT APPLICATION NUMBER: US/10/467,657
 / CURRENT FILING DATE: 2003-08-11
 / PRIOR APPLICATION NUMBER: GB-0103424.8
 / PRIOR FILING DATE: 2001-02-12
 / NUMBER OF SEQ ID NOS: 9218
 / SOFTWARE: SeqWIn99, version 1.04
 / SEQ ID NO: 6780
 / LENGTH: 512
 / TYPE: PRT
 / ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-6780

Query Match 34.4%; Score 37.5; DB 6; Length 512;
 Best Local Similarity 31.8%; Pred. No. 94;
 Matches 7; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

Qy 2 AILIP-ALAGLQVVARAGDIS 22
 ::::|:|:|:|:|
 Db 280 SVIVDLAATGNCETLTPGELS 301

Search completed: January 23, 2006, 12:09:41
 Job time : 4.21397 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:20:59 ; Search time 76.3319 Seconds
(without alignments)
132.392 Million cell updates/sec

Title: US-10-501-838a-11

Sequence: 1 RAFLPLALAGLCQVAPADISS 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_21.*
1: geneseq219808:*
2: geneseq219908:*
3: geneseq220008:*
4: geneseq220018:*
5: geneseq220028:*
6: geneseq220038:*
7: geneseq220048:*
8: geneseq220058:*
9: geneseq220068:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	23	ADBI6896	Adb16896 Rhizobium
2	109	100.0	23	AEBO8228	Aeb08228 Mesorhizo
3	52	47.7	305	AAYS8038	Aay58038 Rice chit
4	51	46.8	37	ADJ36272	Adj36272 Self-coal
5	51	46.8	38	ADJ36281	Adj36281 Self-coal
6	51	46.8	47	ADJ36270	Adj36270 Self-coal
7	51	46.8	49	ADJ36297	Adj36297 Self-coal
8	51	46.8	52	ADJ36285	Adj36285 Self-coal
9	51	46.8	67	ADJ36278	Adj36278 Self-coal
10	51	46.8	132	ADJ36263	Adj36263 Self-coal
11	51	46.8	133	ADJ36265	Adj36265 Self-coal
12	51	46.8	169	ADJ36251	Adj36251 Self-coal
13	51	46.8	341	ADZ47402	Adz47402 341 amino
14	51	46.8	377	ABG19449	Abg19449 Novel hum
15	49	45.0	27	ADJ36132	Adj36132 Self-coal
16	49	45.0	217	ABO61272	Ab061272 Klebsiell
17	48.5	44.5	30	ADBI6921	Adbi6921 E.coli pe
18	48.5	44.5	39	ADBI6903	Adbi6903 E.coli pe
19	48	44.0	22	ADBI6910	Adbi6910 Escherich
20	48	44.0	22	AEBO8244	Aeb08244 Escherich
21	48	44.0	23	ADBI6888	Adbi6888 Escherich
22	48	44.0	23	ADBI6909	Adbi6909 Escherich
23	48	44.0	23	AEBO8220	Aeb08220 Escherich
24	48	44.0	23	AEBO8243	Aeb08243 Escherich

25	48	44.0	24	6	ADBI6908	Adbi6908 Escherich
26	48	44.0	24	9	AEBO8242	Aeb08242 Escherich
27	48	44.0	25	6	ADBI6922	Adbi6922 Escherich
28	48	44.0	25	6	ADBI6906	Adbi6906 Escherich
29	48	44.0	26	6	ADBI6904	Adbi6904 E.coli pe
30	48	44.0	29	6	ADBI6920	Adbi6920 Escherich
31	48	44.0	29	9	AEBO8254	Aeb08254 Penetrati
32	48	44.0	30	6	ADBI6919	Adbi6919 Escherich
33	48	44.0	30	6	ADBI6905	Adbi6905 Escherich
34	48	44.0	30	9	AEBO8239	Aeb08239 Penetrati
35	48	44.0	30	9	AEBO8253	Aeb08253 Penetrati
36	48	44.0	31	6	ADBI6918	Adbi6918 Escherich
37	48	44.0	31	6	AEBO8252	Aeb08252 Penetrati
38	48	44.0	209	6	ADA00709	Ada00709 Modified
39	48	44.0	213	9	AEBO8263	Aeb08263 Escherich
40	48	44.0	258	6	ADA27293	Ada27293 Plasmid P
41	48	44.0	258	6	ADA00702	Ada00702 Modified
42	48	44.0	359	9	ADW95406	Adw95406 Amino aci
43	48	44.0	400	6	ADA27291	Ada27291 Plasmid P
44	48	44.0	400	6	ADA27292	Ada27292 Plasmid P
45	48	44.0	400	6	ADA00700	Ada00700 Modified

ALIGNMENTS

RESULT 1

ID ADBI6896 standard; peptide; 23 AA.

AC ADBI6896;

DT 20-NOV-2003 (first entry)

DE Rhizobium loti MRA4119 penetrating peptide 11.

KW Penetrating peptide; epithelial; endothelial; tight junction; diabetes;

KW Infertility; hormone; vitamin deficiency; neurodegenerative;

KW Cardiovascular; haematological; endocrine disorder; obesity;

KW Neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;

KW Osteopathic; cytostatic; nootropic.

OS Mesorhizobium loti.

PN WO2003066859-A2.

PP 07-FEB-2003; 2003WO-IB000968.

PR 07-FEB-2002; 2002US-0355396P.

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PI Ben-Sasson SA, Cohen E;

DR WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating

PS or preventing e.g. endocrine disorders.

XX Claim 2; Page 14; 60pp; English.

CC This invention relates to a novel peptide sequences capable of

translocating across a biological barrier. Furthermore, it refers to

methods that use these peptides to facilitate penetration of a

biologically active effector molecule such as a drug or other therapeutic

agent across biological barriers e.g. epithelial or endothelial cells

sealed by tight junctions. This peptide is derived from a bacterial

toxigen, an integral membrane or extracellular protein and can comprise an

anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin

or enzyme. The effector molecule, however, can comprise for example

insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
CC cyostatic or neurotropic activities. This peptide is from MMR4119 of
CC Rhizobium loti and is penetrating peptide 11 of the invention.

XX Sequence 23 AA;

XX Query Match 100.0%; Score 109; DB 6; Length 23;

XX Best Local Similarity 100.0%; Pred. No. 1.4e-09;

XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALILPLALAGLCQVARAGDISS 23
Db 1 RALILPLALAGLCQVARAGDISS 23

RESULT 2

AEBO8228 standard; peptide; 23 AA.

AC AEB08228;

DT 25-AUG-2005 (first entry)

DE Mesorhizobium loti MMR4119 penetrating peptide 11, SEQ ID NO: 11.

XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
XX antidiabetic; endocrine disease; gastrointestinal disease;
XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;
XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
XX dementia; multiple sclerosis; immune disorder; Huntington's chorea;
XX anticonvulsant; genetic disorder; cardiovascular disease;
XX cardiovascular-gen.; atherosclerosis; arteriosclerosis;
XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
XX genitourinary disease; hematological disease; antianemic; anemia;
XX autoimmune disease; immunosuppressive; immune deficiency;
XX immunostimulant; infectious disease; antimicrobial; infection;
XX erectile dysfunction; andrology; major depressive disorder;
XX antidepressant; psychiatric disorder; pain; analgesic;
XX bacterial infection; antibacterial; viral infection; virocidic;
XX fungal infection; fungicide; parasitic infection; antiparasitic;
XX renal failure; antifertility; antipneumatic; cytostatic;
XX antiinflammatory; hepatotropic; hepatitis B virus infection.

XX Mesorhizobium loti.

OS US2005136103-A1.

PN 23-JUN-2005.

PD 16-SEP-2004; 2004US-00942300.

PR 17-SEP-2003; 2003US-00664989.

PR 17-SEP-2003; 2003US-00665184.

PR 17-SEP-2003; 2003US-0503615P.

PA (BENS/) BEN-SASSON S A.

PI (COHEN/) COHEN E.

PI Ben-Sasson SA, Cohen E;

PT WPI, 2005-444089/45.

PT Composition used for translocating effectors across barrier such as
PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
XX
XX Claim 46; SEQ ID NO 11; 59pp; English.

CC The present invention relates to a pharmaceutical composition of
CC penetrating peptides for transmembrane delivery of effector. The
CC invention comprises the effector sequentially coupled with a counter ion
CC and at least one hydrophobic agent, where the effector is selectively
CC encapsulated into a complex. The invention is useful for translocating
CC effectors across a biological barrier such as epithelial cells and
CC endothelial cells during treatment and prevention of disease or
CC pathological conditions (including endocrine disorders, diabetes,
CC infertility, hormone deficiencies, osteoporosis, ophthalmological
CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
CC Parkinson's disease, multiple sclerosis, Huntington's disease,
CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
CC coagulable states, coronary disease, cerebrovascular events, metabolic
CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC hematological disorders, anemia of different entities, immunologic and
CC rheumatological disorders, autoimmune diseases, immune deficiencies,
CC infectious diseases, viral infections, bacterial infections, fungal
CC infections, parasitic infections, neoplastic diseases, multi-factorial
CC disorders, impotence, chronic pain, depression, different fibrosis states
CC and short stature) and for mucosal vaccination against anthrax and
CC hepatitis B. The present sequence is the Mesorhizobium loti MMR4119
CC penetrating peptide. This sequence is used in the effective translocation
CC of aminoglycoside antibiotics and antifungal agents across an epithelial
CC barrier.

XX Sequence 23 AA;

XX Query Match 100.0%; Score 109; DB 9; Length 23;

XX Best Local Similarity 100.0%; Pred. No. 1.4e-09;

XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALILPLALAGLCQVARAGDISS 23
Db 1 RALILPLALAGLCQVARAGDISS 23

RESULT 3

AAVS8038 standard; protein; 305 AA.

AC AAVS8038;

DT 06-MAR-2000 (first entry)

DE Rice chitinase protein sequence.

XX Rice; chitinase; pathogen resistance; crop; pathogenic microbe.

OS Oryza sativa.

PN JP11299488-A.

PD 02-NOV-1999.

PD 20-APR-1998; 98JP-00123905.

PR 20-APR-1998; 98JP-00123905.

PA (NORO) NORINSUISANSO SHOKUHN SGO.

PA (SRIB-) SRIBUTSUKI TOKUTSI SANGYO GIUTSU.

DR WPI; 2000-065672/06.

DR N-PSDB; AAZ47465.

PT New rice chitinase cDNA - has lytic activity on filamentous fungi and
PT bacteria.

PT Claim 2; Page 7-8; 10pp; Japanese.

CC This sequence is the rice chitinase of the invention, which has lytic
 CC activity on filamentous fungi and bacteria. A plasmid containing the DNA
 CC and a transformant containing the plasmid can be used for the development
 CC of a recombinant crop, resistant to pathogenic microbes, especially
 CC bacteria and fungi

XX Sequence 305 AA;

Query Match 47.7%; Score 52; DB 3; Length 305;
 Best Local Similarity 55.0%; Pred. No. 15;
 Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 ILIPLALAGLCQVARAGDIS 22
 ID :|||:|||||:
 DB 14 IMVVVALAGLAGTAGTRAGDIA 33

RESULT 4
 ADJ36272
 ID ADJ36272 standard; peptide; 37 AA.

XX ADJ36272;

XX 22-APR-2004 (first entry)

XX Self-coalescing element (SCE) peptide #127.

XX Self-coalescing element; SCE;

XX granulocyte-macrophage colony-stimulating factor; GM-CSF;

XX haematopoietic condition; myelosuppressive disorder; AIDS;

XX infectious disease; autoimmune diseases; allergy; viral infection;

XX bacterial infection; parasitic infection; cancer; melanoma; leukaemia;

XX lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;

XX antibacterial; virucide; antiparasitic; human.

XX Homo sapiens.

XX WO2003102187-A1.

XX 11-DEC-2003.

XX 30-MAY-2003; 2003WO-AU000667.

XX 31-MAY-2002; 2002US-0384878P.

XX (SCEG-) SCEGEN PTY LTD.

XX Koentgen F;

XX WPI; 2004-035460/03.

XX Example 23; SEQ ID NO 221; 219pp; English.

XX The invention relates to an isolated or purified higher order aggregate

XX comprising several chimeric molecules, where each chimeric molecule has a

XX self-coalescing element (SCE) which is fused, linked or associated with a

XX molecule of interest, and where each SCE is capable of causing individual

XX chimeric molecules to coalesce with other chimeric molecules into higher

XX order aggregates under conditions favourable to aggregation. The

XX aggregate is useful for treating or preventing a disease or condition in

XX a patient and has a range of applications including therapeutic,

XX prophylactic and chemical processes. The aggregate comprises a

XX therapeutic polypeptide such as granulocyte-macrophage colony-stimulating

XX factor (GM-CSF), useful for treating haematopoietic conditions such as

XX myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune

XX diseases, allergies, viral, bacterial and parasitic infections and

XX cancers such as melanoma, leukaemia and lymphoma. This sequence

XX represents an SCE peptide of the invention.

SQ Sequence 37 AA;

Query Match 46.8%; Score 51; DB 8; Length 37;
 Best Local Similarity 59.1%; Pred. No. 2.2;
 Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 ALIPLALAGLCQVARAGDIS 23
 ID :|||:|||||:
 DB 4 ALAIIVALAGFATVVAQAGSSS 25

RESULT 5
 ADJ36281
 ID ADJ36281 standard; peptide; 38 AA.

XX ADJ36281;

XX 22-APR-2004 (first entry)

XX Self-coalescing element (SCE) peptide #136.

XX Self-coalescing element; SCE;

XX granulocyte-macrophage colony-stimulating factor; GM-CSF;

XX haematopoietic condition; myelosuppressive disorder; AIDS;

XX infectious disease; autoimmune diseases; allergy; viral infection;

XX bacterial infection; parasitic infection; cancer; melanoma; leukaemia;

XX lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;

XX antibacterial; virucide; antiparasitic; human.

XX Homo sapiens.

XX WO2003102187-A1.

XX 11-DEC-2003.

XX 30-MAY-2003; 2003WO-AU000667.

XX 31-MAY-2002; 2002US-0384878P.

XX (SCEG-) SCEGEN PTY LTD.

XX Koentgen F;

XX WPI; 2004-035460/03.

XX Example 31; SEQ ID NO 230; 219pp; English.

XX The invention relates to an isolated or purified higher order aggregate

XX comprising several chimeric molecules, where each chimeric molecule has a

XX self-coalescing element (SCE) which is fused, linked or associated with a

XX molecule of interest, and where each SCE is capable of causing individual

XX chimeric molecules to coalesce with other chimeric molecules into higher

XX order aggregates under conditions favourable to aggregation. The

XX aggregate is useful for treating or preventing a disease or condition in

XX a patient and has a range of applications including therapeutic,

XX prophylactic and chemical processes. The aggregate comprises a

XX therapeutic polypeptide such as granulocyte-macrophage colony-stimulating

XX factor (GM-CSF), useful for treating haematopoietic conditions such as

XX myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune

XX diseases, allergies, viral, bacterial and parasitic infections and

XX cancers such as melanoma, leukaemia and lymphoma. This sequence

XX represents an SCE peptide of the invention.

QY 2 AIIPLALAGCOVARAGDIS 23
 ||| : ||| ||| : ||| |||
 DB 4 AIAIIVALAGPATVVAQAGSGSS 25

RESULT 6
 ADJ36270
 ID ADJ36270 standard; protein; 47 AA.

ADJ36270;

22-APR-2004 (first entry)

Self-coalescing element (SCE) peptide #125.

Self-coalescing element; SCE;
 granulocyte-macrophage colony-stimulating factor; GM-CSF;
 haematopoietic condition; myelosuppressive disorder; AIDS;
 infectious disease; autoimmune diseases; allergy; viral infection;
 bacterial infection; parasitic infection; cancer; melanoma; leukaemia;
 lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;
 antibacterial; virucide; antiparasitic; human.

Homo sapiens.

MO2003102187-A1.

11-DEC-2003.

30-MAY-2003; 2003WO-AU000667.

31-MAY-2002; 2002US-0384878P.

(SCEG-) SCEGEN PTY LTD.

Koentgen F;

WPI; 2004-035460/03.

Isolated or purified higher order aggregate comprising several chimeric molecules having self-coalescing element associated or fused with molecule, useful in chemical, therapeutic and prophylactic applications.

Example 21; SEQ ID NO 219; 219pp; English.

The invention relates to an isolated or purified higher order aggregate comprising several chimeric molecules, where each chimeric molecule has a self-coalescing element (SCE) which is fused, linked or associated with a molecule of interest, and where each SCE is capable of causing individual chimeric molecules to coalesce with other chimeric molecules into higher order aggregates under conditions favourable to aggregation. The aggregate is useful for treating or preventing a disease or condition in a patient and has a range of applications including therapeutic, prophylactic and chemical processes. The aggregate comprises a therapeutic polypeptide such as granulocyte-macrophage colony-stimulating factor (GM-CSF), useful for treating haematopoietic conditions such as myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune diseases, allergies, viral, bacterial and parasitic infections and cancers such as melanoma, leukaemia and lymphoma. This sequence represents an SCE peptide of the invention.

Sequence 47 AA;

Query Match 46.8%; Score 51; DB 8; Length 47;
 Best Local Similarity 59.1%; Pred. No. 2.8;
 Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIIPLALAGCOVARAGDIS 23
 ||| : ||| ||| : ||| |||
 DB 4 AIAIIVALAGPATVVAQAGSGSS 25

RESULT 7

ADJ36297
 ID ADJ36297 standard; protein; 49 AA.

ADJ36297;

22-APR-2004 (first entry)

Self-coalescing element (SCE) peptide #152.

Self-coalescing element; SCE;
 granulocyte-macrophage colony-stimulating factor; GM-CSF;
 haematopoietic condition; myelosuppressive disorder; AIDS;
 infectious disease; autoimmune diseases; allergy; viral infection;
 bacterial infection; parasitic infection; cancer; melanoma; leukaemia;
 lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;
 antibacterial; virucide; antiparasitic; HIV.

Human immunodeficiency virus.

MO2003102187-A1.

11-DEC-2003.

30-MAY-2003; 2003WO-AU000667.

31-MAY-2002; 2002US-0384878P.

(SCEG-) SCEGEN PTY LTD.

Koentgen F;

WPI; 2004-035460/03.

Isolated or purified higher order aggregate comprising several chimeric molecules having self-coalescing element associated or fused with molecule, useful in chemical, therapeutic and prophylactic applications.

Example 47; SEQ ID NO 246; 219pp; English.

The invention relates to an isolated or purified higher order aggregate comprising several chimeric molecules, where each chimeric molecule has a self-coalescing element (SCE) which is fused, linked or associated with a molecule of interest, and where each SCE is capable of causing individual chimeric molecules to coalesce with other chimeric molecules into higher order aggregates under conditions favourable to aggregation. The aggregate is useful for treating or preventing a disease or condition in a patient and has a range of applications including therapeutic, prophylactic and chemical processes. The aggregate comprises a therapeutic polypeptide such as granulocyte-macrophage colony-stimulating factor (GM-CSF), useful for treating haematopoietic conditions such as myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune diseases, allergies, viral, bacterial and parasitic infections and cancers such as melanoma, leukaemia and lymphoma. This sequence represents an SCE peptide of the invention.

Sequence 49 AA;

Query Match 46.8%; Score 51; DB 8; Length 49;
 Best Local Similarity 59.1%; Pred. No. 3;
 Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIIPLALAGCOVARAGDIS 23
 ||| : ||| ||| : ||| |||
 DB 4 AIAIIVALAGPATVVAQAGSGSS 25

RESULT 8
 ADJ36285
 ID ADJ36285 standard; protein; 52 AA.

ADJ36285;

22-APR-2004 (first entry)

XX Self-coalescing element (SCB) peptide #140.
 XX
 XX Self-coalescing element; SCB;
 XX granulocyte-macrophage colony-stimulating factor; GM-CSF;
 XX haematopoietic condition; myelosuppressive disorder; AIDS;
 XX infectious disease; autoimmune diseases; allergy; viral infection;
 XX bacterial infection; parasitic infection; cancer; melanoma; leukaemia;
 XX lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;
 XX antibacterial; virucide; antiparasitic; human.
 XX
 OS Homo sapiens.
 XX
 XX MO2003102187-A1.
 XX
 XX 11-DEC-2003.
 XX
 XX 30-MAY-2003; 2003MO-AU000667.
 XX
 XX 31-MAY-2002; 2002US-0384878P.
 XX
 XX (SCBG-) SCBGEN PTY LTD.
 XX
 XX Koentgen F;
 XX
 XX WPI; 2004-035460/03.
 XX
 XX Isolated or purified higher order aggregate comprising several chimeric
 PT molecules having self-coalescing element associated or fused with
 PT molecule, useful in chemical, therapeutic and prophylactic applications.
 XX
 XX Example 35; SEQ ID NO 234; 219pp; English.
 XX
 CC The invention relates to an isolated or purified higher order aggregate
 CC comprising several chimeric molecules, where each chimeric molecule has a
 CC self-coalescing element (SCB) which is fused, linked or associated with a
 CC molecule of interest, and where each SCB is capable of causing individual
 CC chimeric molecules to coalesce with other chimeric molecules into higher
 CC order aggregates under conditions favourable to aggregation. The
 CC aggregate is useful for treating or preventing a disease or condition in
 CC a patient and has a range of applications including therapeutic,
 CC prophylactic and chemical processes. The aggregate comprises a
 CC therapeutic polypeptide such as granulocyte-macrophage colony-stimulating
 CC factor (GM-CSF), useful for treating haematopoietic conditions such as
 CC myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune
 CC diseases, allergies, viral, bacterial and parasitic infections and
 CC cancers such as melanoma, leukaemia and lymphoma. This sequence
 CC represents an SCB peptide of the invention.
 CC
 XX
 XX Sequence 52 AA;
 SQ
 Query Match 46.8%; Score 51; DB 8; Length 52;
 Best Local Similarity 59.1%; Pred. No. 3.2;
 Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 OY 2 AIIIPALAGLCQVARAGDIS 23
 |||:|||||:|||||
 Db 4 AIAIAVALAGFATVAQAAGSGSS 25
 |||:|||||:|||||
 RESULT 9
 ADJ36278
 ID ADJ36278 standard; protein; 67 AA.
 XX
 AC ADJ36278;
 XX
 DT 22-APR-2004 (first entry)
 XX
 XX Self-coalescing element (SCB) peptide #133.
 DE
 XX Self-coalescing element; SCB;
 XX granulocyte-macrophage colony-stimulating factor; GM-CSF;
 XX haematopoietic condition; myelosuppressive disorder; AIDS;
 XX

KW infectious disease; autoimmune diseases; allergy; viral infection;
 KW bacterial infection; parasitic infection; cancer; melanoma; leukaemia;
 KW lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;
 KW antibacterial; virucide; antiparasitic; mouse.
 XX
 OS Mus sp.
 XX
 XX MO2003102187-A1.
 XX
 XX 11-DEC-2003.
 XX
 XX 30-MAY-2003; 2003MO-AU000667.
 XX
 XX 31-MAY-2002; 2002US-0384878P.
 XX
 XX (SCBG-) SCBGEN PTY LTD.
 XX
 XX Koentgen F;
 XX
 XX WPI; 2004-035460/03.
 XX
 XX Isolated or purified higher order aggregate comprising several chimeric
 PT molecules having self-coalescing element associated or fused with
 PT molecule, useful in chemical, therapeutic and prophylactic applications.
 XX
 XX Example 28; SEQ ID NO 227; 219pp; English.
 XX
 CC The invention relates to an isolated or purified higher order aggregate
 CC comprising several chimeric molecules, where each chimeric molecule has a
 CC self-coalescing element (SCB) which is fused, linked or associated with a
 CC molecule of interest, and where each SCB is capable of causing individual
 CC chimeric molecules to coalesce with other chimeric molecules into higher
 CC order aggregates under conditions favourable to aggregation. The
 CC aggregate is useful for treating or preventing a disease or condition in
 CC a patient and has a range of applications including therapeutic,
 CC prophylactic and chemical processes. The aggregate comprises a
 CC therapeutic polypeptide such as granulocyte-macrophage colony-stimulating
 CC factor (GM-CSF), useful for treating haematopoietic conditions such as
 CC myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune
 CC diseases, allergies, viral, bacterial and parasitic infections and
 CC cancers such as melanoma, leukaemia and lymphoma. This sequence
 CC represents an SCB peptide of the invention.
 CC
 XX
 XX Sequence 67 AA;
 SQ
 Query Match 46.8%; Score 51; DB 8; Length 67;
 Best Local Similarity 59.1%; Pred. No. 4.2;
 Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 OY 2 AIIIPALAGLCQVARAGDIS 23
 |||:|||||:|||||
 Db 4 AIAIAVALAGFATVAQAAGSGSS 25
 |||:|||||:|||||
 RESULT 10
 ADJ36263
 ID ADJ36263 standard; protein; 132 AA.
 XX
 AC ADJ36263;
 XX
 DT 22-APR-2004 (first entry)
 XX
 XX Self-coalescing element (SCB) peptide #119.
 DE
 XX Self-coalescing element; SCB;
 XX granulocyte-macrophage colony-stimulating factor; GM-CSF;
 XX haematopoietic condition; myelosuppressive disorder; AIDS;
 XX infectious disease; autoimmune diseases; allergy; viral infection;
 XX bacterial infection; parasitic infection; cancer; melanoma; leukaemia;
 XX lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;
 XX antibacterial; virucide; antiparasitic; mouse.
 XX
 OS Mus sp.

```

XX XX WO2003102187-A1.
XX XX
XX PD 11-DEC-2003.
XX XX
XX PF 30-MAY-2003; 2003WO-AU000667.
XX XX
XX PR 31-MAY-2002; 2002US-0384878P.
XX XX
XX PA (SCEG-) SCEGEN PTY LTD.
XX XX
XX PI Koentgen F;
XX XX
XX DR WPI; 2004-035460/03.
XX DR N-PSDB; ADJ36262.
XX XX
XX PT Isolated or purified higher order aggregate comprising several chimeric
XX PT molecules having self-coalescing element associated or fused with
XX PT molecule, useful in chemical, therapeutic and prophylactic applications.
XX XX
XX PS Example 15; SEQ ID NO 212; 219pp; English.
XX XX
XX CC The invention relates to an isolated or purified higher order aggregate
XX CC comprising several chimeric molecules, where each chimeric molecule has a
XX CC self-coalescing element (SCE) which is fused, linked or associated with a
XX CC molecule of interest, and where each SCE is capable of causing individual
XX CC chimeric molecules to coalesce with other chimeric molecules into higher
XX CC order aggregates under conditions favourable to aggregation. The
XX CC aggregate is useful for treating or preventing a disease or condition in
XX CC a patient and has a range of applications including therapeutic,
XX CC prophylactic and chemical processes. The aggregate comprises a
XX CC therapeutic polypeptide such as granulocyte-macrophage colony-stimulating
XX CC factor (GM-CSF), useful for treating haematopoietic conditions such as
XX CC myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune
XX CC diseases, allergies, viral, bacterial and parasitic infections and
XX CC cancers such as melanoma, leukaemia and lymphoma. This sequence
XX CC represents an SCE peptide of the invention.
XX XX
XX SQ Sequence 132 AA;
XX XX
XX Query Match 46.8%; Score 51; DB 8; Length 132;
XX Best Local Similarity 59.1%; Pred. No. 8.7;
XX Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
XX XX
OY 2 AILPLALAGLCQVARAGDIS 23
XX ||| :||| ||| |||
XX Db 5 AIAIAVLAAGPATVAQAGSGSS 26
XX
XX RESULT 11
XX ID ADJ36265 standard; protein; 133 AA.
XX XX
XX AC ADJ36265;
XX XX
XX DT 22-APR-2004 (first entry)
XX XX
XX DE Self-coalescing element (SCE) peptide #120.
XX XX
XX KW Self-coalescing element; SCE;
XX KW granulocyte-macrophage colony-stimulating factor; GM-CSF;
XX KW haematopoietic condition; myelosuppressive disorder; AIDS;
XX KW infectious diseases; autoimmune diseases; allergy; viral infection;
XX KW bacterial infection; parasitic infection; cancer; melanoma; leukaemia;
XX KW lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;
XX KW antibacterial; virucide; antiparasitic; human.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO2003102187-A1.
XX XX
XX PD 11-DEC-2003.
XX XX

```

```

XX PF 30-MAY-2003; 2003WO-AU000667.
XX XX
XX PR 31-MAY-2002; 2002US-0384878P.
XX XX
XX PA (SCEG-) SCEGEN PTY LTD.
XX XX
XX PI Koentgen F;
XX XX
XX DR WPI; 2004-035460/03.
XX DR N-PSDB; ADJ36264.
XX XX
XX PT Isolated or purified higher order aggregate comprising several chimeric
XX PT molecules having self-coalescing element associated or fused with
XX PT molecule, useful in chemical, therapeutic and prophylactic applications.
XX XX
XX PS Example 16; SEQ ID NO 214; 219pp; English.
XX XX
XX CC The invention relates to an isolated or purified higher order aggregate
XX CC comprising several chimeric molecules, where each chimeric molecule has a
XX CC self-coalescing element (SCE) which is fused, linked or associated with a
XX CC molecule of interest, and where each SCE is capable of causing individual
XX CC chimeric molecules to coalesce with other chimeric molecules into higher
XX CC order aggregates under conditions favourable to aggregation. The
XX CC aggregate is useful for treating or preventing a disease or condition in
XX CC a patient and has a range of applications including therapeutic,
XX CC prophylactic and chemical processes. The aggregate comprises a
XX CC therapeutic polypeptide such as granulocyte-macrophage colony-stimulating
XX CC factor (GM-CSF), useful for treating haematopoietic conditions such as
XX CC myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune
XX CC diseases, allergies, viral, bacterial and parasitic infections and
XX CC cancers such as melanoma, leukaemia and lymphoma. This sequence
XX CC represents an SCE peptide of the invention.
XX XX
XX SQ Sequence 133 AA;
XX XX
XX Query Match 46.8%; Score 51; DB 8; Length 133;
XX Best Local Similarity 59.1%; Pred. No. 8.7;
XX Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
XX XX
OY 2 AILPLALAGLCQVARAGDIS 23
XX ||| :||| ||| |||
XX Db 5 AIAIAVLAAGPATVAQAGSGSS 26
XX
XX RESULT 12
XX ID ADJ36251 standard; protein; 169 AA.
XX XX
XX AC ADJ36251;
XX XX
XX DT 22-APR-2004 (first entry)
XX XX
XX DE Self-coalescing element (SCE) peptide #113.
XX XX
XX KW Self-coalescing element; SCE;
XX KW granulocyte-macrophage colony-stimulating factor; GM-CSF;
XX KW haematopoietic condition; myelosuppressive disorder; AIDS;
XX KW infectious diseases; autoimmune diseases; allergy; viral infection;
XX KW bacterial infection; parasitic infection; cancer; melanoma; leukaemia;
XX KW lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;
XX KW antibacterial; virucide; antiparasitic; human.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO2003102187-A1.
XX XX
XX PD 11-DEC-2003.
XX XX
XX PF 30-MAY-2003; 2003WO-AU000667.
XX XX
XX PR 31-MAY-2002; 2002US-0384878P.
XX XX
XX PA (SCEG-) SCEGEN PTY LTD.
XX XX

```

XX Koentgen F;
PI WPI, 2004-035460/03.
DR N-PSDB; AD247402.
XX Isolated or purified higher order aggregate comprising several chimeric
PT molecules having self-coalescing element associated or fused with
PT molecule, useful in chemical, therapeutic and prophylactic applications.
XX Example 10; SEQ ID NO 200; 219pp; English.
XX The invention relates to an isolated or purified higher order aggregate
CC comprising several chimeric molecules, where each chimeric molecule has a
CC self-coalescing element (SCE) which is fused, linked or associated with a
CC molecule of interest, and where each SCE is capable of causing individual
CC chimeric molecules to coalesce with other chimeric molecules into higher
CC order aggregates under conditions favourable to aggregation. The
CC aggregate is useful for treating or preventing a disease or condition in
CC a patient and has a range of applications including therapeutic, a
CC prophylactic and chemical processes. The aggregate comprises a
CC therapeutic polypeptide such as granulocyte-macrophage colony-stimulating
CC factor (GM-CSF), useful for treating haematopoietic conditions such as
CC myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune
CC diseases, allergies, viral, bacterial and parasitic infections and
CC cancers such as melanoma, leukaemia and lymphoma. This sequence
CC represents an SCE peptide of the invention.
XX
SQ Sequence 169 AA;
XX
Query Match 46.8%; Score 51; DB 8; Length 169;
Best Local Similarity 59.1%; Pred. No. 11;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 2 AIIIPALAGLCQVARAGDIS 23
DB 5 AIIAIVALLAGFATVAQAGSGSS 26
XX
RESULT 13
AD247402
ID AD247402 standard; protein; 341 AA.
XX
AC AD247402;
XX
DT 30-JUN-2005 (first entry)
XX
DE 341 amino acid alcohol dehydrogenase protein.
XX
KM alcohol dehydrogenase; enzyme.
XX
OS Gordonia sp.
XX JP2005102511-A.
XX
PD 21-Apr-2005.
XX
PF 26-SEP-2003; 2003JP-00336235.
XX
PR 26-SEP-2003; 2003JP-00336235.
XX
PA (KANF) KANEKA CORP.
XX
DR WPI, 2005-300043/31.
DR N-PSDB; AD247403.
XX
PT New secondary alcohol dehydrogenase of type Adh1, Adh2 or Adh3, useful
PT for producing optically active secondary alcohols.
XX
PS Claim 7; SEQ ID NO 1; 20pp; Japanese.
XX
CC The invention relates to a novel secondary alcohol dehydrogenase, of the
CC type Adh1. The alcohol dehydrogenase being capable of oxidizing 2-

CC propanol by using NAD+ as a coenzyme and producing acetone, having a
CC molecular weight of 67000 as determined by a gel filtration technique,
CC having an optimum temperature of 30 deg C at the time of oxidation of 2-
CC propanol, and having a fully defined 341 amino acid sequence (AD247402),
CC given in the specification. The invention further comprises: a DNA
CC encoding the alcohol dehydrogenase, and having a sequence chosen from a
CC fully defined 1026 nucleotide (AD247403), 1056 nucleotide (AD247407), or
CC 1524 nucleotide (AD247413) given in the specification, or a sequence of
CC SEQ ID NO. 2, 4 or 6 with one or more base deletion, insertion,
CC substitution or additions; a recombinant vector containing the alcohol
CC dehydrogenase DNA; a transformed organism containing the vector; and a
CC method for producing the alcohol dehydrogenase. The alcohol dehydrogenase
CC is useful for producing optically active secondary alcohols. This
CC sequence represents the 341 amino acid alcohol dehydrogenase protein of
CC the invention.
XX
SQ Sequence 341 AA;
XX
Query Match 46.8%; Score 51; DB 9; Length 341;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 1 RAIIIPALAGLCQVARAGD 20
DB 85 KTIHLPLITCGLCRACRDED 104
XX
RESULT 14
ABG19449
ID ABG19449 standard; protein; 377 AA.
XX
AC ABG19449;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19440.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSB-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
PI
XX WPI, 2001-639362/73.
DR N-PSDB; AAS83636.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 49808; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 377 AA;

Query Match 46.8%; Score 51; DB 4; Length 377;
 Best Local Similarity 71.4%; Pred. No. 27;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 6 PLALAGLCQVAPAG 19
 |||||:
 17 PLGLAGLCRADAG 30

RESULT 15

ID ADJ36132 standard; peptide; 27 AA.

AC ADJ36132;

DT 22-APR-2004 (first entry)

DE Self-coalescing element (SCE) peptide #68.

XX Self-coalescing element; SCE;
 KW granulocyte-macrophage colony-stimulating factor; GM-CSF;
 KW haematopoietic condition; myelosuppressive disorder; AIDS;
 KW infectious disease; autoimmune diseases; allergy; viral infection;
 KW bacterial infection; parasitic infection; cancer; melanoma; leukaemia;
 KW lymphoma; cytostatic; anti-HIV; immunosuppressive; antiallergic;
 KW antibacterial; virucide; antiparasitic.

OS Bacillus subtilis.

FN WO2003102187-A1.

PD 11-DEC-2003.

PF 30-MAY-2003; 2003WO-AU000667.

PR 31-MAY-2002; 2002US-0384878P.

PA (SCEG-) SCEGEN PTY LTD.

PI Koentgen F;

DR WPI; 2004-035460/03.

XX Isolated or purified higher order aggregate comprising several chimeric
 PT molecules having self-coalescing element associated or fused with
 PT molecule, useful in chemical, therapeutic and prophylactic applications.

PS Claim 9; SEQ ID NO 79; 219pp; English.

XX The invention relates to an isolated or purified higher order aggregate
 CC comprising several chimeric molecules, where each chimeric molecule has a
 CC self-coalescing element (SCE) which is fused, linked or associated with a
 CC molecule of interest, and where each SCE is capable of causing individual
 CC chimeric molecules to coalesce with other chimeric molecules into higher
 CC order aggregates under conditions favourable to aggregation. The
 CC aggregate is useful for treating or preventing a disease or condition in
 CC a patient and has a range of applications including therapeutic.

CC prophylactic and chemical processes. The aggregate comprises a
 CC therapeutic polypeptide such as granulocyte-macrophage colony-stimulating
 CC factor (GM-CSF), useful for treating haematopoietic conditions such as
 CC myelosuppressive disorders e.g. AIDS, infectious diseases, autoimmune
 CC diseases, allergies, viral, bacterial and parasitic infections and
 CC cancers such as melanoma, leukaemia and lymphoma. This sequence
 CC represents an SCE peptide of the invention.

XX SQ Sequence 27 AA;

Query Match 45.0%; Score 49; DB 8; Length 27;
 Best Local Similarity 59.1%; Pred. No. 3.2;
 Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 2 AILPLALAGLCQVAPAGDIS 23
 |||||:
 5 AIAIAVALAGFATVAGAGIPSS 26

Search completed: January 23, 2006, 10:28:26
 Job time : 77.3319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:31:19 ; Search time 73.3188 Seconds
(without alignments)
221.323 Million cell updates/sec

Title: US-10-501-838a-11
Perfect score: 109
Sequence: 1 RALPLALAGLCQVARGDIS 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: uniprot_05.80: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	197	2	Q98ER3_RHILO
2	56	51.4	305	2	Q49827_ORYSA
3	56	51.4	341	2	Q82NB0_STRAW
4	54	49.5	341	2	Q8CK64_STRCO
5	54	49.5	454	2	Q4UY24_XANCP
6	54	49.5	454	2	Q8P5Z3_XANCP
7	52	47.7	305	2	Q04139_ORYSA
8	51	46.8	167	2	Q9KTX9_STRCO
9	51	46.8	341	2	Q768S9_PACTO
10	51	46.8	863	2	Q5Y0T1_GALPH
11	50	45.9	222	2	Q5L6Z3_CHLAB
12	50	45.9	328	2	Q8ZX81_PYPAB
13	50	45.9	366	2	Q73T22_MYCPA
14	50	45.9	400	2	Q7NYV4_CHRYO
15	49	45.0	35	2	Q45665_BACSU
16	49	45.0	609	2	Q8DJC0_SYNEL
17	48.5	44.5	454	2	Q4PHI3_USYMA
18	48.5	44.5	661	2	Q53WK7_ORYSA
19	48	44.0	98	2	Q51PJ0_MAGGR
20	48	44.0	207	1	RUYA_CHLCV
21	48	44.0	208	1	Y1637_YERPE
22	48	44.0	208	1	Q66903_YERPS
23	48	44.0	213	1	YCFC_ECO57
24	48	44.0	213	1	YCFC_ECOL6
25	48	44.0	213	1	YCFC_ECOL1
26	48	44.0	213	1	YCFC_SHIFL
27	48	44.0	358	2	Q742G5_MYCPA
28	48	44.0	411	2	DHE2_ARATH
29	48	44.0	411	2	Q941A5_BRANA
30	48	44.0	435	2	Q9VG46_DROME
31	48	44.0	866	2	Q87TA2_VIBPA

ALIGNMENTS

32	47	43.1	300	2	Q4LME1_9NRUK	Q4LME1_burkholderi
33	47	43.1	447	2	Q8KRS3_TRELI	Q8KRS3_thermoccoc
34	47	43.1	447	2	Q97XK2_9EURY	Q97XK2_thermoccoc
35	47	43.1	478	1	NDOU_RHOCA	P50973_rhodobacter
36	47	43.1	653	2	Q74ZU0_ASHGO	Q74ZU0_ashbya_goss
37	47	43.1	926	2	Q9NYAO_HUMAN	Q9NYAO_homo_sapien
38	47	43.1	1056	2	Q6P2R5_HUMAN	Q6P2R5_homo_sapien
39	47	43.1	1058	2	Q9Y4G2_HUMAN	Q9Y4G2_homo_sapien
40	46.5	42.7	502	2	Q5PF20_SALPA	Q5PF20_salmonella
41	46.5	42.7	504	2	Q5YKT4_SALCH	Q5YKT4_salmonella
42	46.5	42.7	506	1	NORR_SALTY	Q8Z4C6_salmonella
43	46.5	42.7	506	1	NORR_SALTY	Q8Z4C6_salmonella
44	46.5	42.7	673	2	Q4UTU1_XANCP	Q4UTU1_xanthomonas
45	46.5	42.7	679	2	Q8P919_XANCP	Q8P919_xanthomonas

RESULT 1
ID Q98ER3_RHILO PRELIMINARY; PRT; 197 AA.
AC Q98ER3;
DT 04-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE M1r4119 protein.
GN OrderedLocusNames=mlr4119;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_Taxid=381;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idehara K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno M.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti (MAFF303099)."
RD DINA:R8307338 (2000).
DR EMBL:BA000012; EMBL:5085441; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 197 AA; 21171 MW; 749EF54576AA99F1 CRC64;
Query Match 100.0%; Score 109; DB 2; Length 197;
Best Local Similarity 100.0%; Prod. No. 4.6e-08;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RALPLALAGLCQVARGDIS 23
Db 2 RALPLALAGLCQVARGDIS 24
RESULT 2
ID Q49827_ORYSA PRELIMINARY; PRT; 305 AA.
AC Q49827;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Chitinase (EC 3.2.1.14).
GN Name=P0679C12.21; Synonyms=P0489B03.31;
OS Oryza sativa (Japanese cultivar-group). Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_Taxid=39947;
RN (1)
RP NUCLEOTIDE SEQUENCE.

RC Tissue=Callus;
 RX MEDLINE=98162724; PubMed=9501993;
 RA Nagasaka H., Yamamoto K., Shomura A., Koga-Ban Y., Takasuga A.,
 RA Yano M., Minobe Y., Sasaki T.;
 RT "Rice class III chitinase homologues isolated by random cloning of
 RT rice cDNAs";
 RL DNA Res. 4:379-385 (1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsunoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayaishi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijiabata S., Honda M., Ichikawa Y., Idozuma A., Iijima M., Ikeda M.,
 RA Ikono M., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saij S., Sakai K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshinara R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1.";
 RL Nature 420:312-316 (2002).
 DR EMBL, D55711: BAA23809.1; -, mRNA.
 DR EMBL, AP003287: BAD82025.1; -, Genomic DNA.
 DR EMBL, AP003794: BAD82632.1; -, Genomic DNA.
 DR PIR, JCS844: JCS844.
 DR HSSP, P23472: 2HWL.
 DR Gramene, O49827: -.
 DR GO, GO:0004568; F:chitinase activity; IEA.
 DR GO, GO:0016798; F:hydrolyase activity; acting on glycosyl bonds; IEA.
 DR GO, GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO, GO:0008152; P:metabolism; IEA.
 DR InterPro, IPR001223: Glyco_hydro_18.
 DR InterPro, IPR001579: Glyco_hydro_18A.
 DR Pfam, PF00704: Glyco_hydro_18; 1.
 DR PROSITE, PS01095: CHITINASE_16; 1.
 KM GlycoSIDase; Hydrolyase.
 SQ SEQUENCE 305 AA; 32202 MW; 50FDS143FC9A87B8 CRC64;
 Query Match 51.4%; Score 56; DB 2; Length 305;
 Best Local Similarity 60.0%; Pred. No. 5.8;
 Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Oy 3 ILIPLAAGLCQVADAGDIS 22
 Db 14 IMVVVLAGLAAGARAGDIA 33
 RESULT 3
 Q82NB0 STRAW PRELIMINARY; PRT; 341 AA.
 ID Q82NB0;
 AC Q82NB0;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative alcohol dehydrogenase.
 GN Name=adhA3; OrderedLocusNames=SAV1393;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OK NCBI_TaxID=33903;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RC MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
 RA Ikeda H., Ishikawa T., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sasaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis";
 RL Nat. Biotechnol. 21:526-531 (2003).

RN [2]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa T., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sasaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- COPFACTOR: Zinc (By similarity).
 DR EMBL, BA000030: BAC69103.1; -, Genomic DNA.
 DR HSSP, Q9Y9P9: 1H2B.
 DR GO, GO:0046872; F:metal ion binding; IEA.
 DR GO, GO:0016491; F:oxidoreductase activity; IEA.
 DR GO, GO:0008270; F:zinc ion binding; IEA.
 DR InterPro, IPR002328; ADH_SF_Zn.
 DR InterPro, IPR002328; ADH_SF_Zn.
 DR Pfam, PF0107: ADH_zinc_N; 1.
 DR PROSITE, PS00059: ADH_ZINC; 1.
 KM Complete proteome; Metal-binding; Oxidoreductase; Zinc.
 SQ SEQUENCE 341 AA; 35570 MW; F364468FDD68DB80 CRC64;
 Query Match 51.4%; Score 56; DB 2; Length 341;
 Best Local Similarity 61.1%; Pred. No. 6.4;
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 Oy 3 ILIPLAAGLCQVADAGD 20
 Db 87 ILHPLVTCGLCRACRAGD 104
 RESULT 4
 Q8CK64 STRCO PRELIMINARY; PRT; 341 AA.
 ID Q8CK64 STRCO PRELIMINARY;
 AC Q8CK64;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative alcohol dehydrogenase (Zinc-binding).
 GN OrderedLocusNames=SCO0259; ORFNames=SCF1.01, SCF20.05;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OK NCBI_TaxID=1902;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=A3(2) / M145;
 RC MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147 (2002).
 CC -1- COPFACTOR: Zinc (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 DR EMBL, AL939104: CAD55432.1; -, Genomic DNA.
 DR PIR, T36399: T36399.
 DR HSSP, Q9Y9P9: 1H2B.
 DR GO, GO:0016491; F:oxidoreductase activity; IEA.
 DR GO, GO:0008270; F:zinc ion binding; IEA.
 DR InterPro, IPR002085; ADH_SF_Zn.
 DR InterPro, IPR002328; ADH_SF_Zn.
 DR InterPro, IPR002328; ADH_SF_Zn.

DR Pfam; PF00107; ADH zinc N; 1.
 DR PRINTS; PRO1415; ANKYRIN.
 DR PROSITE; PS00059; ADH_ZINC; 1.
 KW Complete proteome; Metal-binding; Oxidoreductase; Zinc.
 SQ SEQUENCE 341 AA; 35625 MW; 25454E27A8B2028A CRC64;

Query Match 49.5%; Score 54; DB 2; Length 341;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 PAILPALAGLCQVARRAD 20
 DB 85 KVIHVLITCGLCRACRAGD 104

RESULT 5
 Q4UY24_XANCP PRELIMINARY; PRT; 454 AA.

ID Q4UY24; AC Q4UY24;
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=XC_0975;
 OS Xanthomonas campestris pv. campestris str. 8004.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 NCBI_TaxID=314565;

NUCLEOTIDE SEQUENCE.
 RC STRAIN=8004;
 RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
 Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
 Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
 Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
 He C.-Z.;
 RA Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 RL EMBL; CP000050; AY48049.1; -; Genomic DNA.
 DR Hypothetical protein.
 KW KW
 SQ SEQUENCE 454 AA; 48539 MW; 93367B69F487B6 CRC64;

Query Match 49.5%; Score 54; DB 2; Length 454;
 Best Local Similarity 52.6%; Pred. No. 16;
 Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 5 IPLPALAGLCQVARRADISS 23
 DB 7 LPLAVAGIAGVQAGDLS 25

RESULT 6
 ID Q8P523_XANCP PRELIMINARY; PRT; 454 AA.
 AC Q8P523;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DE Hypothetical protein XCC3189.
 GN OrderedLocuNames=XCC3189;
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 NCBI_TaxID=340;

NUCLEOTIDE SEQUENCE.
 RP STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
 Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
 Camargo L.E.A., Camarote G., Cavanha P., Cardoso J., Chabergo F.,
 Chapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cristiano-Santos J.R.,
 El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,

RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A.J., Sena J.A.D., Silva C.J. de Souza R.P.,
 RA Spindola L.A.P., Takita M.A., Tamura R.E., Teixeira B.C., Tezra R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
 RA Setubal J.C., Kiteajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities";
 RL Nature 417:459-463(2002).
 DR EMBL; AB012436; AA042459.1; -; Genomic DNA.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 454 AA; 48486 MW; C3977A69E9E58A9B CRC64;

Query Match 49.5%; Score 54; DB 2; Length 454;
 Best Local Similarity 52.6%; Pred. No. 16;
 Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 5 IPLPALAGLCQVARRADISS 23
 DB 7 LPLAVAGIAGVQAGDLS 25

RESULT 7
 ID 004139 ORYSA PRELIMINARY; PRT; 305 AA.
 AC 004139;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Basic class III chitinase OsChib3 precursor.
 OS Oryza sativa (japonica cultivar-group).
 OC Bacteria; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzae; Oryza.
 NCBI_TaxID=39947;

NUCLEOTIDE SEQUENCE.
 RP Truong N.H., Itoh Y.;
 RT "Nucleotide sequence of rice chitinase class III cDNA";
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB003195; BAA22266.1; -; mRNA.
 DR HSSP; P23472; ZHW.
 DR Gramene; 004139; -;
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR01223; Glyco_hydro_18.
 DR InterPro; IPR01579; Glyco_hydro_18AS.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR PROSITE; PS01095; CHITINASE_18; 1.
 KW Signal.

FT SIGNAL 1
 SQ SEQUENCE 305 AA; 32261 MW; B2FB58A462P3BCB9 CRC64;

Query Match 47.7%; Score 52; DB 2; Length 305;
 Best Local Similarity 55.0%; Pred. No. 23;
 Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 3 IPLPALAGLCQVARRADISS 22
 DB 14 INVVALAGIAGAGTRAGDIA 33

RESULT 8
 ID Q9KX9_STRCO PRELIMINARY; PRT; 167 AA.
 AC Q9KX9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

```
DE Hypothetical protein SC03001.
GN OrderedLocNames=SC03001; ORNames=SC03.03c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxId=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.B., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Latke U., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939114; CAB90912.1; -; Genomic DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 167 AA; 17358 MW; 5566B0A2EFC6D3B6 CRC64;

Query Match 46.8%; Score 51; DB 2; Length 167;
Best Local Similarity 38.1%; Pred. No. 19;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

OY 3 ILIPLALAGLCQVARAGDIS 23
DB 3 VYVPLTSLGLAEHRAHAGELGT 23

RESULT 9
O768S9_9ACTO PRELIMINARY; PRT; 341 AA.
ID O768S9_9ACTO PRELIMINARY; PRT; 341 AA.
AC O768S9_9ACTO PRELIMINARY; PRT; 341 AA.
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Alcohol dehydrogenase.
GN Name=adh1;
OS Gordonia sp. TV-5.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacteriaceae; Gordoniaceae; Gordonia.
OX NCBI_TaxId=235467;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TV-5;
RX PubMed=14645271; DOI=10.1128/JB.185.24.7120-7128.2003;
RA Kotani T., Yamamoto T., Yurimoto H., Sakai Y., Kato N.;
RT "Propane monooxygenase and NAD+-dependent secondary alcohol
RT dehydrogenase in propane metabolism by Gordonia sp. strain TV-5.";
RL J. Bacteriol. 185:7120-7128(2003).
CC -1- COPACTOR: zinc (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DR EMBL; AB112920; BAD03962.1; -; Genomic DNA.
DR HSSP; PA0394; IAGN.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0008270; P:zinc ion binding; IEA.
DR InterPro; IPR002085; ADH_SF_Zn.
DR InterPro; IPR002328; ADH_Zn.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00107; ADH_zinc_N; 1.
DR PRINTS; PR01415; ANKYRIN.
DR PROSITE; PS00059; ADH_ZINC_1.
KW Metal-binding; NAD; Oxidoreductase; Zinc
SQ SEQUENCE 341 AA; 35534 MW; D0856552372D28F5 CRC64;

Query Match 46.8%; Score 51; DB 2; Length 341;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 RALILPLALAGLCQVARAGD 20
DB 85 KVIILHPLITCGLCRACRDGD 104

RESULT 10
O5Y0T1_9ALPH PRELIMINARY; PRT; 863 AA.
ID O5Y0T1_9ALPH PRELIMINARY; PRT; 863 AA.
AC O5Y0T1;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Viron glycoprotein H.
GN Name=U12;
OS Cercopithecine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxId=10317;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15629785; DOI=10.1016/j.virol.2004.09.042;
RA Tyler S.D., Peters G.A., Severini A.;
RT "Complete genome sequence of cercopithecine herpesvirus 2 (SA8) and
RT comparison with other simplexviruses.";
RL Virology 331:429-440(2005).
DR EMBL; AY174813; AAD88087.1; -; Genomic DNA.
SQ SEQUENCE 863 AA; 91455 MW; 8D31A653F84B5113 CRC64;

Query Match 46.8%; Score 51; DB 2; Length 863;
Best Local Similarity 64.7%; Pred. No. 79;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 ILIPLALAGLCQVARAG 19
DB 840 VLIAPALAGLARVARTG 856

RESULT 11
O5L6Z3_CHLAB PRELIMINARY; PRT; 222 AA.
ID O5L6Z3_CHLAB PRELIMINARY; PRT; 222 AA.
AC O5L6Z3;
DT 01-FEB-2005 (TReMBLrel. 29, Created)
DT 01-FEB-2005 (TReMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TReMBLrel. 29, Last annotation update)
DE Putative holliday junction DNA helicase RuVA.
GN Name=ruva; OrderedLocNames=CAB119;
OS Chlamydomonada abortus.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomonada.
OX NCBI_TaxId=83555;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=S26/3;
RX PubMed=15837807; DOI=10.1101/gr.3684805;
RA Thomson N.R., Yeates C., Bell K., Holden M.T.G., Bentley S.D.,
RA Livingstone M., Cerdeno-Tarraga A.-M., Harris B., Doggett J.,
RA Ormond D., Mungall K., Clarke K., Feltwell T., Hance Z., Sanders M.,
RA Quail M.A., Price C., Barrall B.G., Parkhill J., Donabottom D.;
RT "The Chlamydomonada abortus genome sequence reveals an array of
RT variable proteins that contribute to interspecies variation.";
RL Genome Res. 15:629-640(2005).
DR EMBL; CR648038; CAH63577.1; -; Genomic DNA.
KM Complete proteome; Helicase.
SQ SEQUENCE 222 AA; 24583 MW; FFE8E37A9311FEF8 CRC64;

Query Match 45.9%; Score 50; DB 2; Length 222;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 ALIPLALAGLCQVARAGDIS 23
DB 104 AILNTFSLGLCQVARAGDIS 125
```

```

RESULT 12
Q82X81_PYPAR
ID Q82X81_PYPAR PRELIMINARY; PRT; 238 AA.
AC Q82X81;
DT 01-MAR-2002 (TRMBLrel. 20, Created)
DT 01-MAR-2002 (TRMBLrel. 20, Last sequence update)
DE Hypothetical protein PA01409.
GN OrderedLocusNames=PA01409;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
  Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
  aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AB009820; AAL63468.1; -; Genomic_DNA.
KM Complete proteome.
SQ SEQUENCE 238 AA; 25783 MW; C93A278583E2610B CRC64;

Query Match 45.9%; Score 50; DB 2; Length 238;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 RALIPALAGCQVABGRIS 22
DB 164 KATLPLWISKGLAAARAGDGLA 185

RESULT 13
Q73T22_MYCPA
ID Q73T22_MYCPA PRELIMINARY; PRT; 366 AA.
AC Q73T22;
DT 05-JUL-2004 (TRMBLrel. 27, Created)
DT 05-JUL-2004 (TRMBLrel. 27, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP3576;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Ammons A., Alt D., Kapur V.;
RL Submitted (SFP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017340; AA806126.1; -; Genomic_DNA.
DR InterPro; IPR005106; Hmer_dh_NAD_bind.
DR Pfam; PF03447; NAD_binding_3; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 366 AA; 38253 MW; 81A91EAB6B72DB3 CRC64;

Query Match 45.9%; Score 50; DB 2; Length 366;
Best Local Similarity 90.9%; Pred. No. 54;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 LIPALAGLCQ 14
DB 151 LIPALAGTCQ 161

RESULT 14
Q7NVV4_CHRVO

```

```

ID Q7NVV4_CHRVO PRELIMINARY; PRT; 400 AA.
AC Q7NVV4;
DT 01-MAR-2004 (TRMBLrel. 26, Created)
DT 01-MAR-2004 (TRMBLrel. 26, Last sequence update)
DE Probable MFS transporter.
GN OrderedLocusNames=CV1168;
OS Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 12472 / DSM 30191;
RX Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimarães C.T.,
  Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
  Alves-Gomes J.A., Andrade E.M., Araújo J., de Araujo M.F.F.,
  Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
  Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
  Bordignon J., Brígido M.M., Brito C.A., Brocchi M., Burley H.A.,
  Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
  Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,
  Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcão C.L.,
  Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
  Ferro M.I.T., Franco G.R., Freitas N.S.A., Furian L.R.,
  RA Gazitapanelli R.T., Gomes E.A., Gonçalves P.R., Grangelito T.B.,
  RA Grattapaglia D., Griesard E.C., Hanna B.S., Jardim S.N., Laurino J.,
  RA Leal L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
  Madeira H.M.T., Manfio G.P., Maranhão A.O., Martins W.S.,
  di Mauro S.M.Z., de Medeiros S.R.B., Meisner R.V., Moreira M.A.M.,
  RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
  RA Paixão R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
  RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
  RA Ramalho-Neto C.B., Reis A.M.M., Rigó L.U., Rondinelli E.,
  RA Santos E.B.P., Santos F.R., Schneider M.P.C., Senauez H.N.,
  RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
  RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.T.,
  RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
  RA Vettore A., Massen R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
  remarkable and exploitable bacterial adaptability."
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR EMBL; AB016914; AA058843.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002086; Aldehyde_dehydrog.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
KM Complete proteome.
SQ SEQUENCE 400 AA; 41898 MW; 340CB45288B9597 CRC64;

Query Match 45.9%; Score 50; DB 2; Length 400;
Best Local Similarity 52.6%; Pred. No. 58;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RALIPALAGCQVABGRIS 19
DB 82 RALIPALAGCQVABGRIS 100

RESULT 15
Q45665_BACSU
ID Q45665_BACSU PRELIMINARY; PRT; 35 AA.
AC Q45665;
DT 01-NOV-1996 (TRMBLrel. 01, Created)
DT 01-NOV-1996 (TRMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRMBLrel. 24, Last annotation update)
DE OmpA protein.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;

```

```
BN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=M168;
RX MEDLINE=87222417; PubMed=3108260;
RA Ikemura H., Takagi H., Inouye M.;
RT "Requirement of pro-sequence for the production of active subtilisin E
   in Escherichia coli.";
RL J. Biol. Chem. 262:7859-7864(1987).
DR EMBL, M16639; AAA22743.1; -; Genomic_DNA.
DR PIR, I39969; I39969.
SQ SEQUENCE 35 AA; 3608 MW; 6B2ECS381FC704B2 CRC64;

Query Match          45.0%; Score 49; DB 2; Length 35;
Best Local Similarity 59.1%; Pred NO; 10;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AILPLALAGLCQVARAGDIS 23
   |||:|||||:|||||
DB 5 AIAIVALLAGPATVAQAQIPSS 26
```

Search completed: January 23, 2006, 10:40:49
Job time : 75.318 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:52:43 ; Search time 52.8297 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838A-11
Perfect score: 109
Sequence: 1 RALPLALAGLCQVARAGDIS 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications, AA, Main:
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	23	4	US-10-665-184-11
2	109	100.0	23	5	US-10-942-300-11
3	109	100.0	23	5	US-10-501-838A-11
4	56	51.4	305	4	US-10-437-963-104241
5	56	51.4	341	4	US-10-156-761-8932
6	51	46.8	37	4	US-10-449-831A-221
7	51	46.8	38	4	US-10-449-831A-230
8	51	46.8	47	4	US-10-449-831A-219
9	51	46.8	49	4	US-10-449-831A-9
10	51	46.8	52	4	US-10-449-831A-234
11	51	46.8	67	4	US-10-449-831A-227
12	51	46.8	132	4	US-10-449-831A-212
13	51	46.8	133	4	US-10-449-831A-214
14	51	46.8	169	4	US-10-449-831A-200
15	51	46.8	377	5	US-10-450-763-49808
16	49	45.0	27	4	US-10-449-831A-79
17	49	45.0	609	5	US-10-732-923-2197
18	48.5	44.5	29	5	US-10-501-838A-18
19	48.5	44.5	30	5	US-10-501-838A-20
20	48.5	44.5	606	4	US-10-437-963-17453
21	48	44.0	22	4	US-10-665-184-27
22	48	44.0	22	5	US-10-942-300-27
23	48	44.0	22	5	US-10-501-838A-27
24	48	44.0	23	4	US-10-665-184-3
25	48	44.0	23	4	US-10-665-184-26
26	48	44.0	23	5	US-10-942-300-3
27	48	44.0	23	5	US-10-942-300-26

ALIGNMENTS

28	48	44.0	23	5	US-10-501-838A-3	Sequence 3, Appl
29	48	44.0	23	5	US-10-501-838A-26	Sequence 26, Appl
30	48	44.0	24	4	US-10-665-184-25	Sequence 25, Appl
31	48	44.0	24	5	US-10-942-300-25	Sequence 25, Appl
32	48	44.0	24	5	US-10-501-838A-25	Sequence 25, Appl
33	48	44.0	25	5	US-10-501-838A-19	Sequence 19, Appl
34	48	44.0	25	5	US-10-501-838A-23	Sequence 23, Appl
35	48	44.0	26	5	US-10-501-838A-21	Sequence 21, Appl
36	48	44.0	29	4	US-10-665-184-37	Sequence 37, Appl
37	48	44.0	29	5	US-10-942-300-37	Sequence 37, Appl
38	48	44.0	29	5	US-10-501-838A-37	Sequence 37, Appl
39	48	44.0	30	4	US-10-665-184-22	Sequence 22, Appl
40	48	44.0	30	4	US-10-665-184-36	Sequence 36, Appl
41	48	44.0	30	5	US-10-942-300-22	Sequence 22, Appl
42	48	44.0	30	5	US-10-942-300-36	Sequence 36, Appl
43	48	44.0	30	5	US-10-501-838A-36	Sequence 36, Appl
44	48	44.0	30	5	US-10-501-838A-36	Sequence 36, Appl
45	48	44.0	31	4	US-10-665-184-35	Sequence 35, Appl

RESULT 1
US-10-665-184-11
; Sequence 11, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Rhizobium loti
US-10-665-184-11

Query Match 100.0%; Score 109; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 RALPLALAGLCQVARAGDIS 23
Db 1 RALPLALAGLCQVARAGDIS 23

RESULT 2
US-10-942-300-11
; Sequence 11, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 11
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Rhizobium loti
US-10-942-300-11

Query Match 100.0%; Score 109; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALIFPLALAGLCQVARAGDISS 23
Db 1 RALIFPLALAGLCQVARAGDISS 23

RESULT 3
US-10-501-838A-11
; Sequence 11, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Rhizobium loti
US-10-501-838A-11

Query Match 100.0%; Score 109; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RALIFPLALAGLCQVARAGDISS 23
Db 1 RALIFPLALAGLCQVARAGDISS 23

RESULT 4
US-10-437-963-104241
; Sequence 104241, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221) B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204366
; SEQ ID NO 104241
; LENGTH: 305
; TYPE: PRT

;; ORGANISM: Oryza sativa
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4530_101594C.1.pep
US-10-437-963-104241

Query Match 51.4%; Score 56; DB 4; Length 305;
Best Local Similarity 60.0%; Pred. No. 1.7;
Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ILIPLALAGLCQVARAGDIS 22
Db 14 IMVVVALAGLAAGARAGDIA 33

RESULT 5
US-10-156-761-8932
; Sequence 8932, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8932
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8932

Query Match 51.4%; Score 56; DB 4; Length 341;
Best Local Similarity 61.1%; Pred. No. 1.9;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 ILIPLALAGLCQVARAGD 20
Db 87 ILHPLVTCGLCRACRAGD 104

RESULT 6
US-10-449-831A-221
; Sequence 221, Application US/10449831A
; Publication No. US20040029179A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; TITLE OF INVENTION: Higher molecular weight entities and uses therefor
; FILE REFERENCE: 2385978
; CURRENT APPLICATION NUMBER: US/10/449,831A
; CURRENT FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US9N 60/384878
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 237
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 221
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human gamma MSH chimeric peptide
US-10-449-831A-221

Query Match 46.8%; Score 51; DB 4; Length 37;

Best Local Similarity 59.1%; Pred. No. 1.1;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIIPLALAGLCOVARAGDISS 23
DB 4 AIAIAVALAGFATVAQAQSGSS 25

RESULT 7

US-10-449-831A-230
Sequence 230, Application US/10449831A

Publication No. US20040029179A1
GENERAL INFORMATION:

APPLICANT: Koentgen, Frank
TITLE OF INVENTION: Higher molecular weight entities and uses therefor

FILE REFERENCE: 2385978
CURRENT APPLICATION NUMBER: US/10/449,831A

CURRENT FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: USSN 60/384878

PRIOR FILING DATE: 2002-05-31
NUMBER OF SEQ ID NOS: 237

SOFTWARE: PatentIn version 3.2
SEQ ID NO 230

LENGTH: 38
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Human IL-2 chimeric peptide I
US-10-449-831A-230

Query Match 46.8%; Score 51; DB 4; Length 38;
Best Local Similarity 59.1%; Pred. No. 1.1;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIIPLALAGLCOVARAGDISS 23
DB 4 AIAIAVALAGFATVAQAQSGSS 25

RESULT 8

US-10-449-831A-219
Sequence 219, Application US/10449831A

Publication No. US20040029179A1
GENERAL INFORMATION:

APPLICANT: Koentgen, Frank
TITLE OF INVENTION: Higher molecular weight entities and uses therefor

FILE REFERENCE: 2385978
CURRENT APPLICATION NUMBER: US/10/449,831A

CURRENT FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: USSN 60/384878

PRIOR FILING DATE: 2002-05-31
NUMBER OF SEQ ID NOS: 237

SOFTWARE: PatentIn version 3.2
SEQ ID NO 219

LENGTH: 47
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Human beta MSH chimeric peptide
US-10-449-831A-219

Query Match 46.8%; Score 51; DB 4; Length 47;
Best Local Similarity 59.1%; Pred. No. 1.4;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIIPLALAGLCOVARAGDISS 23
DB 4 AIAIAVALAGFATVAQAQSGSS 25

RESULT 9
US-10-449-831A-9
Sequence 9, Application US/10449831A

Publication No. US20040029179A1
GENERAL INFORMATION:

APPLICANT: Koentgen, Frank
TITLE OF INVENTION: Higher molecular weight entities and uses therefor

FILE REFERENCE: 2385978
CURRENT APPLICATION NUMBER: US/10/449,831A

CURRENT FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: USSN 60/384878

PRIOR FILING DATE: 2002-05-31
NUMBER OF SEQ ID NOS: 237

SOFTWARE: PatentIn version 3.2
SEQ ID NO 9

LENGTH: 49
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: HIV gp120 chimeric peptide II
US-10-449-831A-9

Query Match 46.8%; Score 51; DB 4; Length 49;
Best Local Similarity 59.1%; Pred. No. 1.5;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIIPLALAGLCOVARAGDISS 23
DB 4 AIAIAVALAGFATVAQAQSGSS 25

RESULT 10

US-10-449-831A-234
Sequence 234, Application US/10449831A

Publication No. US20040029179A1
GENERAL INFORMATION:

APPLICANT: Koentgen, Frank
TITLE OF INVENTION: Higher molecular weight entities and uses therefor

FILE REFERENCE: 2385978
CURRENT APPLICATION NUMBER: US/10/449,831A

CURRENT FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: USSN 60/384878

PRIOR FILING DATE: 2002-05-31
NUMBER OF SEQ ID NOS: 237

SOFTWARE: PatentIn version 3.2
SEQ ID NO 234

LENGTH: 52
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Human TNF-alpha chimeric peptide II
US-10-449-831A-234

Query Match 46.8%; Score 51; DB 4; Length 52;
Best Local Similarity 59.1%; Pred. No. 1.6;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 AIIPLALAGLCOVARAGDISS 23
DB 4 AIAIAVALAGFATVAQAQSGSS 25

RESULT 11

US-10-449-831A-227
Sequence 227, Application US/10449831A

Publication No. US20040029179A1
GENERAL INFORMATION:

APPLICANT: Koentgen, Frank
TITLE OF INVENTION: Higher molecular weight entities and uses therefor

FILE REFERENCE: 2385978
CURRENT APPLICATION NUMBER: US/10/449,831A

CURRENT FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: USSN 60/384878

PRIOR FILING DATE: 2002-05-31
NUMBER OF SEQ ID NOS: 237

SOFTWARE: PatentIn version 3.2

```

; SEQ ID NO 227
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Murine GHRH chimeric peptide
US-10-449-831A-227

Query Match          46.8%; Score 51; DB 4; Length 67;
Best Local Similarity 59.1%; Pred. No. 2.1;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy      2 AIIPLPALGICQVAPAGDIS 23
Db      4 AIAIAVALGAFVAVQAQSGSS 25

RESULT 12
US-10-449-831A-212
; Sequence 212, Application US/10449831A
; Publication No. US20040029179A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; TITLE OF INVENTION: Higher molecular weight entities and uses therefor
; FILE REFERENCE: 2385978
; CURRENT APPLICATION NUMBER: US/10/449,831A
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: USSN 60/384878
; NUMBER OF SEQ ID NOS: 237
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 212
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Murine MCP-1 chimeric construct
US-10-449-831A-212

Query Match          46.8%; Score 51; DB 4; Length 132;
Best Local Similarity 59.1%; Pred. No. 4.3;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy      2 AIIPLPALGICQVAPAGDIS 23
Db      5 AIAIAVALGAFVAVQAQSGSS 26

RESULT 13
US-10-449-831A-214
; Sequence 214, Application US/10449831A
; Publication No. US20040029179A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; TITLE OF INVENTION: Higher molecular weight entities and uses therefor
; FILE REFERENCE: 2385978
; CURRENT APPLICATION NUMBER: US/10/449,831A
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: USSN 60/384878
; NUMBER OF SEQ ID NOS: 237
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 214
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human MCP-1 chimeric construct
US-10-449-831A-214

Query Match          46.8%; Score 51; DB 4; Length 133;
Best Local Similarity 59.1%; Pred. No. 4.3;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
```

```

Qy      2 AIIPLPALGICQVAPAGDIS 23
Db      5 AIAIAVALGAFVAVQAQSGSS 26

RESULT 14
US-10-449-831A-200
; Sequence 200, Application US/10449831A
; Publication No. US20040029179A1
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; TITLE OF INVENTION: Higher molecular weight entities and uses therefor
; FILE REFERENCE: 2385978
; CURRENT APPLICATION NUMBER: US/10/449,831A
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: USSN 60/384878
; NUMBER OF SEQ ID NOS: 237
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 200
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human IL-2 chimeric construct
US-10-449-831A-200

Query Match          46.8%; Score 51; DB 4; Length 169;
Best Local Similarity 59.1%; Pred. No. 5.6;
Matches 13; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy      2 AIIPLPALGICQVAPAGDIS 23
Db      5 AIAIAVALGAFVAVQAQSGSS 26

RESULT 15
US-10-450-763-49808
; Sequence 49808, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 49808
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (64)..(231)
; OTHER INFORMATION: Ribosomal L18p/L5e family domain identified by Pfam.
; OTHER INFORMATION: accession name Ribosomal_L18p, E-value=0.015, Pfam score of -25.0
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(377)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
```

US-10-450-763-49808

Query Match 46.8%; Score 51; DB 5; Length 377;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 PLALAGLCQVARRG 19
| | | | | | | | | |
Db 17 PLGLAGLCRAADAG 30

Search completed: January 23, 2006, 12:08:58
Job time : 53.8297 secs

THE BOSTON PUBLIC LIBRARY
ASTOR LENOX TILDEN FOUNDATION
125 NASSAU ST. N.Y.C. 4

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:35:50 ; Search time 22.9258 seconds
(without alignments)
90.156 Million cell updates/sec

Title: US-10-501-838A-12
Perfect score: 124
Sequence: 1 MRNLTKTSLLAGLCTAAQWFEVTH 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5/COMB.pep:*
2: /cgn2_6/prodata/1/aa/6/COMB.pep:*
3: /cgn2_6/prodata/1/aa/H/COMB.pep:*
4: /cgn2_6/prodata/1/aa/PCTUS/COMB.pep:*
5: /cgn2_6/prodata/1/aa/RE/COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	38.7	429	US-09-922-501-10	Sequence 10, Appl
2	45	36.3	217	US-09-489-039A-7789	Sequence 7789, Ap
3	44	35.3	580	US-08-677-049-12	Sequence 12, Appl
4	44	35.5	1146	US-09-824-734-2	Sequence 2, Appl
5	43	34.7	26	US-09-348-578-1	Sequence 1, Appl
6	43	34.7	26	US-09-699-684-1	Sequence 1, Appl
7	43	34.7	27	US-09-348-578-2	Sequence 2, Appl
8	43	34.7	27	US-09-699-684-2	Sequence 2, Appl
9	43	34.7	32	US-09-348-578-7	Sequence 7, Appl
10	43	34.7	32	US-09-699-684-7	Sequence 7, Appl
11	43	34.7	69	US-09-149-476-655	Sequence 655, App
12	43	34.7	114	US-09-902-540-10640	Sequence 10640, A
13	43	34.7	461	US-09-489-039A-12176	Sequence 12176, A
14	43	34.7	504	US-09-107-433-4993	Sequence 4993, Ap
15	43	34.7	510	US-08-255-670A-2	Sequence 2, Appl
16	43	34.7	519	US-09-008-271A-9	Sequence 9, Appl
17	43	34.7	519	US-09-968-415-9	Sequence 9, Appl
18	43	34.7	520	US-09-149-476-351	Sequence 351, App
19	43	34.7	628	US-09-252-991A-32414	Sequence 32414, A
20	42.5	34.3	482	US-09-949-016-7966	Sequence 7966, Ap
21	42.5	34.3	659	US-09-902-540-11293	Sequence 11293, A
22	42	33.9	171	US-09-724-864-59	Sequence 59, Appl
23	42	33.9	538	US-08-994-076-2	Sequence 2, Appl
24	42	33.9	538	US-09-643-476-2	Sequence 2, Appl
25	41.5	33.5	586	US-10-104-047-3411	Sequence 3411, Ap
26	41.5	33.5	806	US-09-902-540-12422	Sequence 12422, A
27	41	33.1	203	US-09-328-714A-6	Sequence 6, Appl

28	41	33.1	229	US-09-270-767-44828	Sequence 44828, A
29	41	33.1	356	US-08-887-534A-72	Sequence 72, Appl
30	41	33.1	356	US-09-527-431-72	Sequence 72, Appl
31	41	33.1	356	US-09-446-861-72	Sequence 72, Appl
32	41	33.1	357	US-09-489-039A-10505	Sequence 10505, A
33	41	33.1	363	US-08-956-171E-5237	Sequence 5237, Ap
34	41	33.1	363	US-08-781-968A-5237	Sequence 5237, Ap
35	41	33.1	426	US-09-328-352-6416	Sequence 6416, Ap
36	41	33.1	533	US-09-252-991A-25841	Sequence 25841, A
37	41	33.1	644	US-10-104-047-2945	Sequence 2945, Ap
38	41	33.1	898	US-10-449-315-11	Sequence 11, Appl
39	40.5	32.7	216	US-09-543-681A-7993	Sequence 7993, Ap
40	40	32.3	28	US-09-348-578-3	Sequence 3, Appl
41	40	32.3	28	US-09-699-684-3	Sequence 3, Appl
42	40	32.3	131	US-09-270-767-32294	Sequence 32294, A
43	40	32.3	131	US-09-270-767-47511	Sequence 47511, A
44	40	32.3	175	US-09-489-039A-11431	Sequence 11431, A
45	40	32.3	270	US-09-489-039A-14315	Sequence 14315, A

ALIGNMENTS

```
RESULT 1
US-09-922-501-10
; Sequence 10, Application US/09922501
; Patent No. 6720168
; GENERAL INFORMATION:
; APPLICANT: Dattois, Veronique A.
; APPLICANT: Hoch, James A.
; APPLICANT: Valle, Fernando
; APPLICANT: Kumar, Manoj
; TITLE OF INVENTION: 2, 5-DKG PERMEASES
; FILE REFERENCE: P-SR 4877
; CURRENT APPLICATION NUMBER: US/09/922,501
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/633,294
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 09/677,032
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PaatSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Pantoea citrea
US-09-922-501-10

Query Match      38.7%  Score 48; DB 2; Length 429;
Best Local Similarity 61.1%; Pred. No. 12;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Cy      1 MRNLTKTSLLAGLCTAA 18
Db      79 VRNLVFTSLILWGFCAAA 96

RESULT 2
US-09-489-039A-7789
; Sequence 7789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7789
; LENGTH: 217
```

;; TYPE: PRT
;; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match 36.3%; Score 45; DB 2; Length 217;
Best Local Similarity 45.0%; Pred. No. 17;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 2 RLNTKTSLLAGLCTAAQMV 21
DB 7 KNYDYITLALGVCGQARLV 26

RESULT 3
US-08-677-049-12
; Sequence 12, Application US/08677049
; Patent No. 5858707
; GENERAL INFORMATION:
; APPLICANT: Guimaraes, M. Jorge
; APPLICANT: Bazan, J. Fernando
; APPLICANT: McClanahan, Terrill K.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: PURIFIED MAMMALIAN NUCLEOBASE PERMEASES;
; TITLE OF INVENTION: NUCLEIC ACIDS; ANTIBODIES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,049
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,788
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-852-9196
; TELEFAX: 415-496-1200
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 207..236
; OTHER INFORMATION: /note="Encompasses TM 4 of Figure
; OTHER INFORMATION: 4"
; NAME/KEY: Region
; LOCATION: 437..499
; OTHER INFORMATION: /note="Encompasses TM 9 and TM 10
; OTHER INFORMATION: of Figure 4"
US-08-677-049-12

Query Match 36.3%; Score 45; DB 1; Length 580;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 4 LTKTSLLAGLCTAAQMV 20
DB 110 LVSTSLVSGLSAQM 126

RESULT 4
US-09-824-734-2
; Sequence 2, Application US/09824734
; Patent No. 6727408
; GENERAL INFORMATION:
; APPLICANT: ZHU, JIAN-KANG
; APPLICANT: SHI, HUAZHONG
; APPLICANT: ISHITANI, MANABU
; APPLICANT: STEVENSON, BECKY
; TITLE OF INVENTION: PROTEINS AND DNA RELATED TO SALT TOLERANCE IN PLANTS
; FILE REFERENCE: 205644US20
; CURRENT APPLICATION NUMBER: US/09/824,734
; CURRENT FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: US 60/194,648
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1146
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-824-734-2

Query Match 35.5%; Score 44; DB 2; Length 1146;
Best Local Similarity 61.1%; Pred. No. 1,7e+02;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 6 KTSLLAGLCTAAQMV 23
DB 159 KTSLLGLSLATDPVAV 176

RESULT 5
US-09-348-578-1
; Sequence 1, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(26)
; OTHER INFORMATION: Description of Artificial Sequence: Oppa secretion signal
US-09-348-578-1

Query Match 34.7%; Score 43; DB 2; Length 26;
Best Local Similarity 52.9%; Pred. No. 3.1;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 RLNTKTSLLAGLCTA 17
DB 1 MTNITKSLVAAQVLA 17

RESULT 6
US-09-699-684-1
Sequence 1, Application US/09699684
Patent No. 6436674
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/699,684
CURRENT FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 26
TYPE: PRT
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(26)
OTHER INFORMATION: Description of Artificial Sequence:Opa secretion signal
US-09-699-684-1

Query Match 34.7%; Score 43; DB 2; Length 26;
Best Local Similarity 52.9%; Pred. No. 3.1;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 MENTKTSLLAGLCTA 17
DB 1 MTNITKRSLLVAAGVLA 17

RESULT 7
US-09-348-578-2
Sequence 2, Application US/09348578
Patent No. 6160089
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/348,578
CURRENT FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: JP 193003/1998
EARLIER FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 27
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Modified Opa secretion
OTHER INFORMATION: signal
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(27)
US-09-348-578-2

Query Match 34.7%; Score 43; DB 2; Length 27;
Best Local Similarity 55.0%; Pred. No. 3.2;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MENTKTSLLAGLCTAQM 20
DB 1 MTNITKRSLLVAAGVLAALM 20

RESULT 8
US-09-699-684-2
Sequence 2, Application US/09699684
Patent No. 6436674
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/699,684
CURRENT FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 27
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Modified Opa secretion
OTHER INFORMATION: signal
NAME/KEY: SIGNAL
LOCATION: (1)..(27)
US-09-699-684-2

Query Match 34.7%; Score 43; DB 2; Length 27;
Best Local Similarity 55.0%; Pred. No. 3.2;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 1 MENTKTSLLAGLCTAQM 20
DB 1 MTNITKRSLLVAAGVLAALM 20

RESULT 9
US-09-348-578-7
Sequence 7, Application US/09348578
Patent No. 6160089
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/348,578
CURRENT FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: JP 193003/1998
EARLIER FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 7
LENGTH: 32
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Modified Opa secretion
OTHER INFORMATION: signal
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(32)
US-09-348-578-7

```
Query Match          34.7% Score 43; DB 2; Length 32;
Best Local Similarity 61.1%; Pred. No. 3.9;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Cy 1 MRNLTKTSLLAGLCTAA 18
   ||:|||||
Db 1 MTNITKRSLLLLLVAA 18

RESULT 10
US-09-699-684-7
; Sequence 7, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified Opa secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)...(32)
US-09-699-684-7

Query Match          34.7% Score 43; DB 2; Length 32;
Best Local Similarity 61.1%; Pred. No. 3.9;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Cy 1 MRNLTKTSLLAGLCTAA 18
   ||:|||||
Db 1 MTNITKRSLLLLLVAA 18

RESULT 11
US-09-149-476-655
; Sequence 655, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
```

```
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
```


EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 34.7%; Score 43; DB 2; Length 69;
Best Local Similarity 62.5%; Pred. No. 9.5;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 12 AGCTAAGV--FVTH 25
DB 10 AGCTAAGVKEFVTH 25

RESULT 12
US-09-902-540-10640
Sequence 10640, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 10640
LENGTH: 114
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-10640

Query Match 34.7%; Score 43; DB 2; Length 114;
Best Local Similarity 33.3%; Pred. No. 17;
Matches 7; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRUTKTSLLAGCTAAGV 21
DB 4 LRPLATRAVYVLAGCAVARI 24

RESULT 13
US-09-489-039A-12176

Sequence 12176, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12176
LENGTH: 461
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12176

Query Match 34.7%; Score 43; DB 2; Length 461;
Best Local Similarity 50.0%; Pred. No. 87;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Cy 6 KTSLLAGICTAAQNV 21
|:|:|:|:|:|:
Db 334 KNTLLIAGWVAARI 349

RESULT 14
US-09-107-433-4993
Sequence 4993, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4993:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae

FEATURE:
NAME/KEY: mlec_feature
LOCATION: (B) LOCATION 1...504
SEQUENCE DESCRIPTION: SEQ ID NO: 4993:
US-09-107-433-4993

Query Match 34.7%; Score 43; DB 1; Length 510;
Best Local Similarity 55.6%; Pred. No. 98;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Cy 4 LTKTSLLAGICTAAQNV 21
|:|:|:|:|:|:
Db 15 LKTFVLMVGLATVAFMV 32

RESULT 15
US-08-255-670A-2
Sequence 2, Application US/08255670A
Patent No. 5691180
GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B
TITLE OF INVENTION: N-ACETYL-GALACTOSAMINE-TRANSFERASE CDNA
SEQUENCE AND EXPRESSION PRODUCTS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/255,670A
FILING DATE: 09-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5691180man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2363-092-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TRIM: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 510 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-255-670A-2

Query Match 34.7%; Score 43; DB 1; Length 510;
Best Local Similarity 55.6%; Pred. No. 98;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Cy 4 LTKTSLLAGICTAAQNV 21
|:|:|:|:|:|:
Db 15 LKTFVLMVGLATVAFMV 32

Search completed: January 23, 2006, 10:46:47
Job time : 22.9258 secs

ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/183,914
FILING DATE: 19-JULY-2005
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/008,271
FILING DATE: 16-Jan-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: ENDCNOT03
CLONE: 2170967
US-11-183-914-9

Query Match 34.7%; Score 43; DB 7; Length 519;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 12 AGLCTAAQMV--FVTH 25
DB 459 AGACTAAAFKKEFVTH 474

RESULT 3
US-10-485-517-125
Sequence 125, Application US/10485517
Publication No. US20050256299A1
GENERAL INFORMATION:
APPLICANT: University of Sheffield
APPLICANT: Biosynexus Incorporated
APPLICANT: Foster, Simon
TITLE OF INVENTION: Antigenic Polypeptides
FILE REFERENCE: P100629WC
CURRENT APPLICATION NUMBER: US/10/485,517
CURRENT FILING DATE: 2004-02-02
PRIOR APPLICATION NUMBER: GB 0118825.9
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: GB 0200349.9
PRIOR FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 424
SOFTWARE: PatentIn version 3.1
SEQ ID NO 125
LENGTH: 361
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-485-517-125

Query Match 33.1%; Score 41; DB 6; Length 361;

Best Local Similarity 55.6%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRNLTKTSLLAGLCTAA 18
DB 8 MRLTKTSLALGLTTGA 25

RESULT 4
US-10-485-517-295
Sequence 295, Application US/10485517
Publication No. US20050256299A1
GENERAL INFORMATION:
APPLICANT: University of Sheffield
APPLICANT: Biosynexus Incorporated
APPLICANT: Foster, Simon
TITLE OF INVENTION: Antigenic Polypeptides
FILE REFERENCE: P100629WC
CURRENT APPLICATION NUMBER: US/10/485,517
CURRENT FILING DATE: 2004-02-02
PRIOR APPLICATION NUMBER: GB 0118825.9
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: GB 0200349.9
PRIOR FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 424
SOFTWARE: PatentIn version 3.1
SEQ ID NO 295
LENGTH: 361
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-485-517-295

Query Match 33.1%; Score 41; DB 6; Length 361;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRNLTKTSLLAGLCTAA 18
DB 8 MRLTKTSLALGLTTGA 25

RESULT 5
US-11-145-631-11
Sequence 11, Application US/11145631
Publication No. US20060003409A1
GENERAL INFORMATION:
APPLICANT: Berlin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-329001
CURRENT APPLICATION NUMBER: US/11/145,631
CURRENT FILING DATE: 2005-06-06
PRIOR APPLICATION NUMBER: US/09/841,739
PRIOR FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: US 09/697,089
PRIOR FILING DATE: 2000-10-26
PRIOR APPLICATION NUMBER: US 60/161,822
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 898
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: majority sequence
US-11-145-631-11

Query Match 33.1%; Score 41; DB 7; Length 898;
Best Local Similarity 46.7%; Pred. No. 55;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRNLTKTSLLAGLCT 15

DB 189 LQNLQKTPLEFVAIC 203

RESULT 6

US-11-000-463-824
Sequence 824, Application US/11000463
Publication No. US20050266423A1
GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Chen, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhiwei
APPLICANT: Wehman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Drmanac, Radolje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIPACN
CURRENT APPLICATION NUMBER: US/11/000,463
PRIOR FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 824
LENGTH: 168
TYPE: PRT
ORGANISM: Homo sapiens
US-11-000-463-824

Query Match 32.3%; Score 40; DB 7; Length 168;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DB 42 TVSLLLMAGLCT 53

RESULT 7

US-10-485-517-126
Sequence 126, Application US/10485517
Publication No. US2005025629A1
GENERAL INFORMATION:
APPLICANT: University of Sheffield
APPLICANT: Biosynex Incorporated
APPLICANT: Foster, Simon
APPLICANT: Mond, James
TITLE OF INVENTION: Antigenic Polypeptides
FILE REFERENCE: P10629WO
CURRENT APPLICATION NUMBER: US/10/485,517
CURRENT FILING DATE: 2004-02-02
PRIOR APPLICATION NUMBER: GB 0118825.9
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: GB 0200349.9
PRIOR FILING DATE: 2002-01-09
NUMBER OF SEQ ID NOS: 424

SOFTWARE: PatentIn version 3.1
SEQ ID NO 126
LENGTH: 251
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-10-485-517-126

Query Match 32.3%; Score 40; DB 6; Length 251;
Best Local Similarity 56.2%; Pred. No. 19;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

DB 23 MNLTKSLLAGLCT 16

RESULT 8

US-10-510-386-102
Sequence 102, Application US/10510386
Publication No. US2005024922A1
GENERAL INFORMATION:
APPLICANT: Andersen, Jens Tonne
APPLICANT: Clausen, Ib Groth
APPLICANT: Jorgensen, Steen Troels
APPLICANT: Olsen, Peter Bjarke
APPLICANT: Rasmussen, Michael Dolberg
TITLE OF INVENTION: Improved Bacillus Host Cell
FILE REFERENCE: 10294.204-US
CURRENT APPLICATION NUMBER: US/10/510,386
CURRENT FILING DATE: 2004-10-04
NUMBER OF SEQ ID NOS: 248
SOFTWARE: PatentIn version 3.3
SEQ ID NO 102
LENGTH: 524
TYPE: PRT
ORGANISM: Bacillus licheniformis
US-10-510-386-102

Query Match 32.3%; Score 40; DB 6; Length 524;
Best Local Similarity 36.4%; Pred. No. 43;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

DB 3 NUTKSLLAGLCTAQMVEVT 24

US-10-821-234-1114
Sequence 1114, Application US/10821234
Publication No. US2005025511A1
GENERAL INFORMATION:

APPLICANT: Labat, Ivan
APPLICANT: Stache-Crahan, Birgit
APPLICANT: Andarmann, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PC SEQ_genes Version 1.0
SEQ ID NO 1114
LENGTH: 182
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1114

Query Match 31.9%; Score 39.5; DB 6; Length 182;
Best Local Similarity 68.4%; Pred. No. 16;
Matches 13; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 MENTSLAGLCTAQA 19
Db 32 MRLATFLLA-LSTAQA 49

RESULT 10

US-10-524-647-6
; Sequence 6, Application US/10524647
; Publication No. US20050281909A1
; GENERAL INFORMATION:
; APPLICANT: Flachmann, Ralf
; APPLICANT: Sauer, Marc
; APPLICANT: Schopfer, Christel R.
; APPLICANT: Klebsattel, Martin
; APPLICANT: Pfeiffer, Angelika-Maria
; APPLICANT: Luck, Thomas
; APPLICANT: Voeste, Dirk
; TITLE OF INVENTION: Use of astaxanthin-containing plants or parts of plants of the
; FILE REFERENCE: 13173-00004-US
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/009109
; PRIOR FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: DE 102 38 980.2
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 38 978.0
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 38 979.9
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 53 112.9
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: DE 102 58 971.2
; PRIOR FILING DATE: 2002-12-16
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Agrobacterium aurantiacum
US-10-524-647-6

Query Match 31.5%; Score 39; DB 6; Length 242;
Best Local Similarity 39.1%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 3 NLTKTSLLAGLCTAQAQWVPTH 25
Db 10 DLTAATSLVSGGIIAAMWALHVVH 32

RESULT 11

US-10-524-647-10
; Sequence 10, Application US/10524647
; Publication No. US20050281909A1
; GENERAL INFORMATION:
; APPLICANT: Flachmann, Ralf
; APPLICANT: Sauer, Marc
; APPLICANT: Schopfer, Christel R.
; APPLICANT: Klebsattel, Martin
; APPLICANT: Pfeiffer, Angelika-Maria
; APPLICANT: Luck, Thomas
; APPLICANT: Voeste, Dirk
; TITLE OF INVENTION: Use of astaxanthin-containing plants or parts of plants of the
; FILE REFERENCE: 13173-00004-US
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/009109
; PRIOR FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: DE 102 38 980.2
; PRIOR FILING DATE: 2002-08-20

; PRIOR APPLICATION NUMBER: DE 102 38 978.0
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 38 979.9
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 53 112.9
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: DE 102 58 971.2
; PRIOR FILING DATE: 2002-12-16
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Paracoccus marcusii
US-10-524-647-10

Query Match 31.5%; Score 39; DB 6; Length 242;
Best Local Similarity 39.1%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 3 NLTKTSLLAGLCTAQAQWVPTH 25
Db 10 DLTAATSLVSGGIIAAMWALHVVH 32

RESULT 12

US-10-524-647-100
; Sequence 100, Application US/10524647
; Publication No. US20050281909A1
; GENERAL INFORMATION:
; APPLICANT: Flachmann, Ralf
; APPLICANT: Sauer, Marc
; APPLICANT: Schopfer, Christel R.
; APPLICANT: Klebsattel, Martin
; APPLICANT: Pfeiffer, Angelika-Maria
; APPLICANT: Luck, Thomas
; APPLICANT: Voeste, Dirk
; TITLE OF INVENTION: Use of astaxanthin-containing plants or parts of plants of the
; FILE REFERENCE: 13173-00004-US
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: PCT/EP2003/009109
; PRIOR FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: DE 102 38 980.2
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 38 978.0
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 38 979.9
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: DE 102 53 112.9
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: DE 102 58 971.2
; PRIOR FILING DATE: 2002-12-16
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 100
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Paracoccus sp. MBIC1143
US-10-524-647-100

Query Match 31.5%; Score 39; DB 6; Length 242;
Best Local Similarity 39.1%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 3 NLTKTSLLAGLCTAQAQWVPTH 25
Db 10 DLTAATSLVSGGIIAAMWALHVVH 32

RESULT 13
US-11-129-143-181

```

Sequence 181, Application US/11129143
Publication No. US2005026518A1
GENERAL INFORMATION:
APPLICANT: BERRY, Alan
APPLICANT: BRETZEL, Warner
APPLICANT: HUMBLIN, Markus
APPLICANT: LOPEZ-ULIBARRI, Rual
APPLICANT: MAYER, Anne F.
APPLICANT: YELISEEV, Alexei A.
TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
FILE REFERENCE: C38435/121966
CURRENT APPLICATION NUMBER: US/11/129,143
CURRENT FILING DATE: 2005-05-13
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 181
LENGTH: 242
TYPE: PRT
ORGANISM: Paracoccus carolinifaciens E-396.
US-11-129-143-181

Query Match      31.5%; Score 39; DB 7; Length 242;
Best Local Similarity 39.1%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      3 NLTKTSLLAGLCTAQMVFVTH 25
      :|||||:|:|:|:|:|:|:|:|:|
DB      10 DLTATSLIVSGGIAMWLAHVH 32

RESULT 14
US-11-070-080-33
Sequence 33, Application US/11070080
Publication No. US20050287625A1
GENERAL INFORMATION:
APPLICANT: Ye, Rick W.
APPLICANT: MILLER, Edward S.
TITLE OF INVENTION: PROCESS FOR EXPRESSION OF FOREIGN GENES IN METHANE METABOLIZING
FILE REFERENCE: CL-2443 US NA
CURRENT APPLICATION NUMBER: US/11/070,080
CURRENT FILING DATE: 2005-03-02
PRIOR APPLICATION NUMBER: US 60/550385
PRIOR FILING DATE: 2004-03-05
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn version 3.3
SEQ ID NO: 33
LENGTH: 242
TYPE: PRT
ORGANISM: Agrobacterium aurantiacum
US-11-070-080-33

Query Match      31.5%; Score 39; DB 7; Length 242;
Best Local Similarity 39.1%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      3 NLTKTSLLAGLCTAQMVFVTH 25
      :|||||:|:|:|:|:|:|:|:|:|
DB      10 DLTATSLIVSGGIAMWLAHVH 32

RESULT 15
US-10-131-826A-330
Sequence 330, Application US/10131826A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvarioff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.

```

```

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C128
CURRENT APPLICATION NUMBER: US/10/131,826A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO: 330
LENGTH: 428
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-826A-330

Query Match      31.5%; Score 39; DB 6; Length 428;
Best Local Similarity 69.2%; Pred. No. 50;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 RNLTSTLLAGL 14
      :|||||:|:|:|:|:|:|:|:|:|
DB      168 RNLTSTLLAGL 180

```

Search completed: January 23, 2006, 12:09:41
Job time : 3.49345 secs

This Page Blank (upside)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:20:59 ; Search time 82.9694 Seconds
(without alignments)
132.392 Million cell updates/sec

Title: US-10-501-838a-12
Sequence: 1 MRNLTKTSLIAGLCTAAQMFVTH 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Geneseqp1808:*
2: Geneseqp1908:*
3: Geneseqp2008:*
4: Geneseqp2008:*
5: Geneseqp2008:*
6: Geneseqp2003a:*
7: Geneseqp2003b:*
8: Geneseqp2004a:*
9: Geneseqp2005a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	100.0	25	ADBI6897	Bacillus
2	124	100.0	25	ABE08229	Bacillus
3	124	100.0	538	ADM66760	Bacillus
4	108	87.1	23	ADBI6911	Bacillus
5	108	87.1	30	ABE08245	Bacillus
6	108	87.1	30	ADBI6915	Bacillus
7	108	87.1	30	ABE08249	Bacillus
8	51	41.1	72	ABBI7993	Human
9	50	40.3	272	ABG09325	Human
10	48	38.7	308	AAE17303	Human
11	48	38.7	414	AAAI5973	Human
12	48	38.7	414	AAAI5981	Human
13	48	38.7	429	AAU77015	Human
14	47	37.9	203	ABE08262	Human
15	47	37.9	337	ADL90082	Human
16	47	37.9	457	AAE17304	Human
17	46	37.1	109	ADL28905	Human
18	46	37.1	109	ADL28905	Human
19	46	37.1	455	AAU64449	Human
20	46	37.1	455	ABM60968	Human
21	46	37.1	499	ABM64775	Human
22	46	37.1	592	ABG12198	Human
23	46	37.1	701	AAE24209	Human
24	45	36.3	115	AAU66603	Human

25	45	36.3	115	6	ABM63122
26	45	36.3	151	4	AAU45209
27	45	36.3	151	4	ABM41728
28	45	36.3	217	7	ABO61272
29	45	36.3	765	4	ABE62833
30	44.5	35.9	348	4	ABE57778
31	44	35.5	205	4	AAU50230
32	44	35.5	205	9	ABE08261
33	44	35.5	205	9	ABE08261
34	44	35.5	350	8	ADL27102
35	44	35.5	494	4	ABE59717
36	44	35.5	577	8	ADL64513
37	44	35.5	744	4	AAE01693
38	44	35.5	744	5	ABG63970
39	44	35.5	744	8	ADL77235
40	44	35.5	957	8	ADL64511
41	44	35.5	1146	5	AAE79302
42	44	35.5	1146	5	ABE79791
43	44	35.5	1146	8	ADL64508
44	44	35.5	1146	8	ADL64510
45	44	35.5	1146	8	ADL64507

ALIGNMENTS

RESULT 1
ADBI6897
ID ADBI6897 standard; peptide; 25 AA.

XX ADBI6897;
XX 20-NOV-2003 (first entry)
XX DB

Bacillus subtilis NprB penetrating peptide 12.
penetrating peptide; epithelial; endothelial; tight junction; diabetes;
infectivity; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
osteopathic; cytostatic; nootropic.

XX Bacillus subtilis.

XX WO2003066859-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003WO-IB000968.

XX 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Ben-Sasson SA, Cohen B;

XX WPI; 2003-697452/66.

XX New penetrating peptide, useful for preparing a composition for treating

XX or preventing e.g. endocrine disorders.

XX Claim 2; Page 14; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
translocating across a biological barrier. Furthermore, it refers to
methods that use these peptides to facilitate penetration of a
biologically active effector molecule such as a drug or other therapeutic
agent across biological barriers e.g. epithelial or endothelial cells
sealed by tight junctions. This peptide is derived from a bacterial
toxin, an integral membrane or extracellular protein and can comprise an
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
or enzyme. The effector molecule, however, can comprise for example
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiatherosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from NprB of *Bacillus*
 CC subtilis and is penetrating peptide 12 of the invention.

XX Sequence 25 AA;

Query Match 100.0%; Score 124; DB 6; Length 25;
 Best Local Similarity 100.0%; Pred. No. 3.8e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRNLTKTSLLAGCTAAQWVFVTH 25
 |||||
 Db 1 MRNLTKTSLLAGCTAAQWVFVTH 25

RESULT 2
 AEB08229
 ID AEB08229 standard; peptide; 25 AA.

AC AEB08229;

XX 25-AUG-2005 (first entry)

DT *Bacillus subtilis* NprB penetrating peptide 12, SEQ ID NO: 12.

XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antihypertensive;
 XX coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; antiviral;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antirheumatic; cyostatic;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.

XX *Bacillus subtilis*.

OS US2005136103-A1.

XX 23-JUN-2005.

PD 16-SEP-2004; 2004US-00942300.

XX 17-SEP-2003; 2003US-00664989.

PR 17-SEP-2003; 2003US-00665184.

PR 17-SEP-2003; 2003US-0503615P.

XX (BENS/) BEN-SASSON S A.

PA (COHE/) COHEN E.

PI Ben-Sasson SA, Cohen E;

XX WPI; 2005-444089/45.

PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX Claim 56; SEQ ID NO 12; 59pp; English.

XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transcellular delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiency, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the *Bacillus subtilis* NprB
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.

XX Sequence 25 AA;

Query Match 100.0%; Score 124; DB 9; Length 25;
 Best Local Similarity 100.0%; Pred. No. 3.8e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRNLTKTSLLAGCTAAQWVFVTH 25
 |||||
 Db 1 MRNLTKTSLLAGCTAAQWVFVTH 25

RESULT 3
 ADM66760
 ID ADM66760 standard; protein; 538 AA.

AC ADM66760;

XX 03-JUN-2004 (first entry)

DT *Bacillus* sp. thermolysin-like protease (TLP) precursor protein 9.

XX thermolysin-like protease; TLP; SI' site; gluten degradation; wheat;
 XX baking industry; beer clarification; brewing; delinting; skin dewooling;
 XX leather; protein hydrolylate production; artificial sweetener; aspartame;
 XX precursor; enzyme.

XX *Bacillus* sp.

OS WO2004011619-A2.

XX 05-FEB-2004.

PD 28-JUL-2003; 2003WO-US023726.

PR 26-JUL-2002; 2002US-0398656P.

XX (STRA-) STRATAGENE.

PA Clark DD, Bramean JC;

XX WPI; 2004-143847/14.

XX New thermolysin-like protease with substrate specificity for a basic or
 PT an acidic amino acid, useful in biological and biomedical research.

PT identifying therapeutic agents and diagnostic markers, or producing
PT artificial sweeteners.
PS Disclosure; SEQ ID NO 53; 82pp; English.
XX
XX
XX The invention relates to a novel thermolysin-like protease (TLP)
CC comprising an S1' site and modified to have a substrate specificity for a
CC basic or an acidic amino acid. The thermolysin-like protease of the
CC invention may be useful in proteolysis applications, biological and
CC biomedical research, identifying therapeutic agents and diagnostic
CC markers, characterizing cells and organisms that have undergone genetic
CC modifications, identifying unknown illnesses, characterizing polypeptides
CC or identifying biological samples. The thermolysin-like protease may also
CC be useful in industrial processes, such as the degradation of gluten from
CC wheat within the baking industry, clarification of beer within the
CC brewing industry, detaching or degrading of skins within the leather
CC industry, preparation of protein hydrolysates or production of artificial
CC sweeteners like aspartame. The current sequence is that of a TLP
CC precursor protein of the invention.
XX
SQ Sequence 538 AA;
Query Match 100.0%; Score 124; DB 8; Length 538;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNRITKSTLLAGLCTPAQMVFVTH 25
DB 1 MNRITKSTLLAGLCTPAQMVFVTH 25
RESULT 4
ADB16911
ID ADB16911 standard; peptide; 23 AA.
XX
XX ADB16911;
XX
XX 20-NOV-2003 (first entry)
XX
XX Bacillus subtilis NprB penetrating peptide 33.
XX
XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
XX inferility; hormone; vitamin deficiency; neurodegenerative;
XX cardiovascular; haematological; endocrine disorder; obesity;
XX neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
XX osteopathic; cytoskeletal; neurotropic.
XX
XX Bacillus subtilis.
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003MO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX
XX WPI, 2003-697452/66.
XX
XX Now penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Claim 2; Page 15; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells

CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
CC cytostatic or neurotropic activities. This peptide is from NprB of Bacillus
CC subtilis and is penetrating peptide 33 of the invention.
XX
SQ Sequence 23 AA;
Query Match 87.1%; Score 108; DB 6; Length 23;
Best Local Similarity 95.7%; Pred. No. 1.1e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNRITKSTLLAGLCTPAQMVFV 23
DB 1 MNRITKSTLLAGLCTPAQMVFV 23
RESULT 5
AEB08245
ID AEB08245 standard; peptide; 23 AA.
XX
XX AEB08245;
XX
XX 25-AUG-2005 (first entry)
XX
XX Bacillus subtilis NprB penetrating peptide 33, SEQ ID NO: 28.
XX
XX
XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
XX antidiabetic; endocrine disease; gastrointestinal disease;
XX metabolic disorder; hormone deficiency; osteoporosis; osteopathy;
XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
XX demerit; multiple sclerosis; immune disorder; Huntington's chorea;
XX cardiovascular; genetic disorder; cardiovascular disease;
XX anticonvulsant; atherosclerosis; antiarteriosclerotic;
XX coronary artery disease; cardiac; vasodilator; obesity; anorectic;
XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
XX genitourinary disease; haematological disease; antihaemic; anemia;
XX autoimmune disease; immunosuppressive; immune deficiency;
XX immunostimulant; infectious disease; antimicrobial; infection;
XX erectile dysfunction; andrology; major depressive disorder;
XX antidepressant; psychiatric disorder; pain; analgesic;
XX bacterial infection; antibacterial; viral infection; virucide;
XX fungal infection; fungicide; parasitic infection; antiparasitic;
XX renal failure; antifertility; antirheumatic; cytoskeletal;
XX antiinflammatory; hepatotropic; hepatitis B virus infection.
XX
XX Bacillus subtilis.
XX
XX US2005136103-A1.
XX
XX 23-JUN-2005.
XX
XX 16-SEP-2004; 2004US-00942300.
XX
XX 17-SEP-2003; 2003US-00664989.
XX
XX 17-SEP-2003; 2003US-00665184.
XX
XX 17-SEP-2003; 2003US-0503615P.
XX
XX (BENS/) BEN-SASSON S A.
XX
XX (COHE/) COHEN E.
XX
XX Ben-Sasson SA, Cohen E;

XX WP1; 2005-444089/45.
 DR
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 28; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transendothelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatological disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Bacillus subtilis NprB
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 CC
 XX
 SQ Sequence 23 AA;
 Query Match 87.1%; Score 108; DB 9; Length 23;
 Best Local Similarity 95.7%; Pred. No. 1.1e-09;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MRNLTKTSLLAGLCTAAQWVFV 23
 Db 1 MRNLTKTSLLAGLCTAAQWVFV 23
 RESULT 6
 ADB16915
 ID ADB16915 standard; peptide; 30 AA.
 AC
 XX ADB16915;
 DT 20-NOV-2003 (first entry)
 DE
 XX Escherichia coli peptide 3 coupled to imaging compound linker, IBW-005.
 XX
 KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; hematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiant; antiarteriosclerotic; osteopathic; cytostatic; nootropic;
 KW imaging linker; penetrating peptide; IBW-005.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site /note= "N-terminal acetyl"
 FT Modified-site 30
 FT Modified-site /note= "C-terminal amide"
 XX
 XX WO2003066859-A2.
 XX
 PD 14-AUG-2003.

XX
 PE 07-FEB-2003; 2003MO-IB000968.
 PE
 XX
 FR 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA
 XX Ben-Sasson SA, Cohen E;
 PI
 XX WP1; 2003-697452/66.
 DR
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PS
 XX Example 3; Page 40; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 CC cytostatic or nootropic activities. This peptide sequence is IBW-005,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 CC
 XX
 SQ Sequence 30 AA;
 Query Match 87.1%; Score 108; DB 6; Length 30;
 Best Local Similarity 95.7%; Pred. No. 1.5e-09;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MRNLTKTSLLAGLCTAAQWVFV 23
 Db 1 MRNLTKTSLLAGLCTAAQWVFV 23
 RESULT 7
 AEB08249
 ID AEB08249 standard; peptide; 30 AA.
 AC
 XX AEB08249;
 DT 25-AUG-2005 (first entry)
 DE
 XX Penetrating peptide SEQ: 32 used in composition for mucosal vaccination.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW nootropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrologic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;

OS	XX	Unidentified.	
XX	XX	Key	Location/Qualifiers
XX	XX	Misc-difference	1 /note= "N-terminal acylated"
XX	XX	Misc-difference	30 /note= "Optionally C-terminal amide, optionally the free amino group of lysine is acylated with a fatty acid"
XX	XX	US2005136103-A1.	
XX	XX	23-JUN-2005.	
XX	XX	16-SEP-2004; 2004US-00942300.	
XX	XX	17-SEP-2003; 2003US-00664989.	
XX	XX	17-SEP-2003; 2003US-00665184.	
XX	XX	17-SEP-2003; 2003US-0503615P.	
XX	XX	(BENS/) BEN-SASSON S A.	
XX	XX	(COHE/) COHEN E.	
XX	XX	Ben-Saeson SA, Cohen E;	
XX	XX	WPI, 2005-444089/45.	
XX	XX	Composition used for translocating effectors across barrier such as	
XX	XX	epithelial cells during treatment of e.g. endocrine disorders comprises	
XX	XX	effector sequentially coupled with counter ion and hydrophobic agent.	
XX	XX	Claim 63; SEQ ID NO 32; 59pp; English.	
XX	XX	The present invention relates to a pharmaceutical composition of	
XX	XX	penetrating peptides for transepithelial delivery of effector. The	
XX	XX	invention comprises the effector sequentially coupled with a counter ion	
XX	XX	and at least one hydrophobic agent, where the effector is selectively	
XX	XX	encapsulated into a complex. The invention is useful for translocating	
XX	XX	effectors across a biological barrier such as epithelial cells and	
XX	XX	endothelial cells during treatment and prevention of disease or	
XX	XX	pathological conditions (including endocrine disorders, diabetes,	
XX	XX	infertility, hormone deficiencies, osteoporosis, ophthalmological	
XX	XX	disorders, neurodegenerative disorders, Alzheimer's disease, dementia,	
XX	XX	Parkinson's disease, multiple sclerosis, Huntington's disease,	
XX	XX	cardiovascular disorder, atherosclerosis, hyper-coagulable states, hypo-	
XX	XX	coagulable states, coronary disease, cerebrovascular events, metabolic	
XX	XX	disorders, obesity, vitamin deficiencies, renal disorders, renal failure,	
XX	XX	hematological disorders, anemia of different entities, immunologic and	
XX	XX	rheumatologic disorders, autoimmune diseases, immune deficiencies,	
XX	XX	infectious diseases, viral infections, bacterial infections, fungal	
XX	XX	infections, parasitic infections, neoplastic diseases, multi-factorial	
XX	XX	disorders, impotence, chronic pain, depression, different fibrosis states	
XX	XX	and short stature) and for mucosal vaccination against anthrax and	
XX	XX	hepatitis B. The present sequence is a penetrating peptide (1BW-005) used	
XX	XX	in the composition for mucosal vaccination using a counter anion and a	
XX	XX	penetrating peptide.	
XX	XX	Sequence 30 Aa;	

ID	ABBI17993	standard; protein; 72 AA
XX	ABBI17993;	
AC	ABBI17993;	
XX		
XX	23-JAN-2002	(first entry)
DE		
XX	Human nervous system related polypeptide SEQ ID NO 6650.	
KM	Human; neurotropic; neuroprotective; cytosolic; dermatological; virucide;	
KM	Immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;	
KM	antiparkinsonian; antischiz; antianaemic; antiarthritic; cancer;	
KM	antihypertensive; hepatotropic; cerebroprotective; antiinflammatory;	
KW	antiallergic; antidiabetic; antituber; anticonvulsant; antifungal;	
KW	antiparastic; cardiac; immune disorder; cardiovascular disorder;	
KW	neurological disease; infection; nephrotropic; gene therapy; vaccine.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200159063-A2.	
XX		
PD	16-AUG-2001.	
XX		
PF	17-JAN-2001;	2001WO-US001334.
XX		
XX	31-JAN-2000;	2000US-0179065P.
PR	04-FEB-2000;	2000US-0180628P.
PR	24-FEB-2000;	2000US-0184664P.
PR	02-MAR-2000;	2000US-0186350P.
PR	16-MAR-2000;	2000US-0189874P.
PR	17-MAR-2000;	2000US-0190076P.
PR	18-APR-2000;	2000US-0198123P.
PR	19-MAY-2000;	2000US-0205157P.
PR	07-JUN-2000;	2000US-0209467P.
PR	28-JUN-2000;	2000US-0214886P.
PR	30-JUN-2000;	2000US-0215135P.
PR	07-JUL-2000;	2000US-0216647P.
PR	07-JUL-2000;	2000US-0216880P.
PR	11-JUL-2000;	2000US-0217487P.
PR	11-JUL-2000;	2000US-0217496P.
PR	14-JUL-2000;	2000US-0218290P.
PR	26-JUL-2000;	2000US-0220963P.
PR	26-JUL-2000;	2000US-0220964P.
PR	14-AUG-2000;	2000US-0224518P.
PR	14-AUG-2000;	2000US-0224519P.
PR	14-AUG-2000;	2000US-0225213P.
PR	14-AUG-2000;	2000US-0225214P.
PR	14-AUG-2000;	2000US-0225266P.
PR	14-AUG-2000;	2000US-0225267P.
PR	14-AUG-2000;	2000US-0225268P.
PR	14-AUG-2000;	2000US-0225270P.
PR	14-AUG-2000;	2000US-0225447P.
PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226279P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226868P.
PR	22-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-02271009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.

XX WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HSEB) HSEB INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI
 XX MPI: 2001-639362/73.
 DR N-PSDB; AAS73512.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostic, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20, SEQ ID NO 39684; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotide are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 272 AA;
 S0
 Query Match 40.3%; Score 50; DB 4; Length 272;
 Best Local Similarity 58.8%; Pred. No. 25;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 5 TKTSLLAGLCTAQMVT 21
 DB 175 TKTSLAIALCTSAKCV 191
 RESULT 10
 AAE17303 standard; protein; 308 AA.
 XX
 AC AAE17303;
 XX
 DT 18-APR-2002 (first entry)
 XX
 DE Human carcinoembryonic antigen protein, sbg251170CEna #1.
 XX
 XX Human; therapy; wound healing disorder; vaccine; cancer; infection;
 KM autoimmune disorder; haemotopoietic disorder; inflammation; arthritis;
 KM Parkinson's disease; Huntington's chorea; schizophrenia; antiarhythmic;
 KM multiple sclerosis; Alzheimer's disease; analgesic; cardiac; asthma;
 KM leukaemia; stroke; AIDS; bone disease; atherosclerosis; brain disorder;
 KM depression; cardiovascular disease; myocardial infarction; renal failure;

KM respiratory disease; liver disorder; Fanconi's syndrome; spleen disorder;
 KM type II diabetes mellitus; skeletal muscle disorder; immunosuppressive;
 KM hyperlipidemia; renal disease; hypoglycaemia; gastrointestinal disease;
 KM neoplastic; cirrhosis; Hodgkin's disease; neuroleptic; antiinflammatory;
 KM haemostatic; vulnery; anticonvulsant; antirheumatic; neuroprotective;
 KM nephrotropic; hypotensive; vasotropic; cytostatic; cerebroprotective;
 KM allergy; carcinoembryonic antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200198342-A1.
 PD 27-DEC-2001.
 XX
 PF 22-JUN-2001; 2001WO-US019929.
 XX
 PR 22-JUN-2000; 2000US-0213156P.
 PR 22-JUN-2000; 2000US-0213161P.
 XX
 PA (SMK) SMITHKLINE BEECHAM CORP.
 PA (SMK) SMITHKLINE BEECHAM PLC.
 PA (GLAXO) GLAXO GROUP LTD.
 XX
 PI Agarwal P, Cogswell JP, Kabnic KS, Lai Y, Martensen SA;
 PI Murdoch PR, Smith RF, Strum JC, Xiang Z, Xie Q, Rizni SK;
 DR MPI: 2002-139783/18.
 DR N-PSDB; AAD27798.
 XX
 PT Novel secreted and membrane-associated polypeptides and polynucleotides
 PT useful for preventing, ameliorating or correcting dysfunction or disease
 PT including diabetes, cancer, hypertension and growth abnormalities.
 XX
 PS Claim 1; Page 100-101; 138pp; English.
 XX
 CC The invention relates to secreted and membrane-associated polypeptides
 CC and polynucleotides. The sequences of the invention are useful in
 CC diagnostic assays for detecting diseases associated with inappropriate
 CC activity or levels of these polynucleotides, and in identifying their
 CC agonists and antagonists that are potentially useful in therapy. The
 CC sequences of the invention are useful as vaccines for inducing
 CC immunological response. The sequences of the invention are useful for
 CC treating cancers, infections, autoimmune disorders, haematopoietic
 CC disorders, wound healing disorders, cholesterol ester storage disease,
 CC inflammation, congenital muscular dystrophy, junctional epidermolysis
 CC bullosa, Parkinson's disease, Huntington's chorea, multiple sclerosis,
 CC viral and bacterial infections, Alzheimer's disease, asthma, arthritis,
 CC allergies, schizophrenia, sbg442445PROA-associated disorders,
 CC septicemia, porphyria, inflammatory bowel disease, transplant rejection,
 CC graft versus host disease, leukaemia, stroke, acute respiratory disease
 CC syndrome, restenosis, brain injury, AIDS, bone diseases, atherosclerosis,
 CC brain disorders including paraneuronal palsy, myotonic dystrophy,
 CC depression, anxiety disorders and sleep disorders, cardiovascular
 CC diseases including congestive heart failure and myocardial infarction,
 CC respiratory diseases including chronic obstructive pulmonary disease,
 CC acute bronchitis and adult respiratory distress syndrome, liver disorders
 CC including hypercholesterolaemia, hypertriglyceridaemia, cirrhosis, viral
 CC and non-viral hepatitis, type II diabetes mellitus, renal disease
 CC including acute and chronic renal failure, glomerulonephritis, Fanconi's
 CC syndrome, cystinuria, skeletal muscle disorders including hypoglycaemia
 CC and tendinitis, gastrointestinal diseases including intestinal
 CC obstruction and tropical sprue, spleen disorders including hyperlipidemia,
 CC Hodgkin's disease and malignant lymphoma, testicular cancer, male
 CC reproductive diseases including low testosterone and male infertility.
 CC The present sequence is human carcinoembryonic antigen
 XX
 S0 Sequence 308 AA;
 QY
 Query Match 38.7%; Score 48; DB 5; Length 308;
 Best Local Similarity 42.1%; Pred. No. 59;
 Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
 QY 6 KTSLLAGLCTAQMVT 24

DT 21-MAY-2002 (first entry)
 XX 2,5-diketo-D-gluconic acid (DKG) permease prmb.
 DE DKG, 2,5-diketo-D-gluconic acid permease; prmb; enzyyme;
 XX 2-keto-L-gulononic acid; 2-KLG; ascorbic acid; 2-keto-reductase.
 KM Pantoea citrea.
 OS WO200212468-A2.
 XX 14-FEB-2002.
 PD 03-AUG-2001; 2001WO-US024507.
 XX 04-AUG-2000; 2000US-00633294.
 XX 29-SEP-2000; 2000US-00677032.
 PR (MIGR-) MICROGENOMICS INC.
 PA (GENW) GENENCOR INT INC.
 XX Darcols VA, Hoch JA, Valle F, Kumar M;
 PI WPI; 2002-217190/27.
 DR N-PSDB; ABK10185.
 XX Novel isolated polypeptide having 2,5-diketo-D-gluconic acid permease
 PT activity, useful for increasing 2-keto-L-gulononic acid bioproduction, and
 PT thus ascorbic acid production.
 PS Claim 44; Fig 1; 90pp; English.
 XX The invention relates to an isolated polypeptide which has 2,5-diketo-D-
 CC gluconic acid (DKG) permease activity namely PE1, PE6, PK1, prma, prmb
 CC and Y1a2. Also included are the nucleic acids encoding the proteins, an
 CC antibody specific for the proteins, and a vector comprising the nucleic
 CC acid. The vector is useful for enhancing 2-keto-L-gulononic acid (2-KLG)
 CC production which involves expressing the polypeptide encoded by the
 CC nucleic acid molecule in a bacterial cell which expresses an enzyme that
 CC catalyzes the conversion of 2,5-DKG to 2-KLG, and enzymes that catalyse
 CC the conversion of glucose to 2,5-DKG. The bacterial cell is deficient in
 CC endogenous 2-keto-reductase activity and is of the genus Pantoea. The
 CC method further involves converting 2-KLG to ascorbic acid. The permease
 CC is useful for enhancing uptake of 2,5-DKG, and also to stop 2,5-DKG
 CC uptake, where the permease is added to a culture medium to compete with
 CC membrane bound permeases for 2,5-DKG. The permease nucleic acid is
 CC expressed in bacterial cells in order to enhance the rate of uptake of
 CC 2,5-DKG by the cells, and as probes or primers to identify and isolate
 CC 2,5-DKG permease homologues from additional species or as templates for
 CC production of mutant permeases. Expression of one or more 2,5-DKG
 CC permeases in bacterial cells significantly increases the production of 2-
 CC keto-L-gulononic acid bioproduction from D-glucose, which lowers the cost
 CC of commercial production of ascorbic acid. The present sequence
 CC represents a 2,5-DKG permease from P. citrea, prmb
 CC XX
 SO Sequence 429 AA;
 Query Match 38.7%; Score 48; DB 5; Length 429;
 Best Local Similarity 61.1%; Pred. No. 86;
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MRNLTSTLLAGCTAA 18
 : |||||
 Db 79 VRKLVFTSLIMGCABA 96
 RESULT 14
 AEB08262
 ID AEB08262 standard; protein; 203 AA.
 XX AEB08262;
 AC
 XX 25-AUG-2005 (first entry)
 DT

XX Pasteurella multocida protein, SEQ ID NO: 60.
 DE
 XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 KM metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KM degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KM neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KM nocotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KM dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KM anticonvulsant; genetic disorder; cardiovascular disease;
 KM cardiovascular-gen.; atherosclerosis; antiatherosclerotic;
 KM coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KM nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KM genitourinary disease; hematological disease; immunodeficiency;
 KM autoimmune disease; immunosuppressive; immune deficiency;
 KM immunostimulant; infectious disease; antimicrobial; infection;
 KM erectile dysfunction; andrology; major depressive disorder;
 KM antidepressant; psychiatric disorder; pain; analgesic;
 KM bacterial infection; antibacterial; viral infection; virologic;
 KM fungal infection; fungicide; parasitic infection; antiparasitic;
 KM renal failure; antifertility; antineumatic; cytostatic;
 KM antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX Pasteurella multocida.
 OS
 XX US2005136103-A1.
 XX 23-JUN-2005.
 PD 16-SEP-2004; 2004US-00942300.
 XX 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-0065184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 XX Ben-Sasson SA, Cohen E;
 PI WPI; 2005-444089/45.
 DR Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Disclosure; SEQ ID NO 60; 59pp; English.
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes;
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological,
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Pasteurella multocida protein
 CC containing penetrating peptide at N-terminal end.
 XX
 SO Sequence 203 AA;

Query Match	37.9%;	Score 47;	DB 9;	Length 203;
Best Local Similarity	47.6%;	Pred. No. 53;		
Matches 10; Conservative	4;	Mismatches 7;	Indels 0;	Gaps 0;

```

QY      1 MRNLTKSLLAGLCTAQM V 21
          | | : | | | : |
DB      1 MANYDYITLALAGVCQAK V 21

```

RESULT 15

ADL90082
ID ADL90082 standard; protein; 337 AA.

AC ADL90082;

DT 17-JUN-2004 (first entry)

Human carcinoma embryonic antigen precursor, SEQ ID 22.

KW Immune response; immunoglobulin; Ig; human;
KW carcinoma embryonic antigen precursor.

Homo sapiens.

PN WO2004027049-A2

PD 01-APR-2004

18-SEP-2003; 2003WO-US030188.

PR 20-SEP-2002; 2002US-0412219P.
1 MAR 2007 0000Z
2 TWO THREE SEVEN
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832

PR 14-MAR-2003; 2003WO-US007995.

PA (ASTR-) ASTRAL INC.

PI Bot A, Wang L, Smith D, Phillips B,

DR WPI; 2004-295415/27.

PT Generating an immune response to an antigen, useful for generating
PT desired T cell responses comprises administering an immunoglobulin having
PT one peptide epitope of the antigen attached to the immunoglobulin.

PS Disclosure; Fig 11; 154pp; English.

CC The present invention relates to a method for generating an immune
CC response to an antigen in a patient. The method comprises administering
CC to the patient an immunoglobulin (Ig) or its portion where the Ig has at
CC least one peptide epitope of the antigen attached to the Ig or its
CC portion and administering the immunoglobulin or its portion in
CC conjunction with a RNA segment. The present sequence is an antigen
CC sequence, used to illustrate the invention.

SQ Sequence 337 AA;

Query Match	Score	DB	Length
37.9%	47	8	337

Best Local Similarity 47.1%; Pred. No. 94;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy	8	SLLAGLCTAAQMVFT	24
		: :	: :
Db	19	SLACGICQASGQIFIT	35

Search completed: January 23, 2006, 10:28:28
Job time : 84.9694 secs

RESULT 2

QNEBOA
Oligopeptide-binding protein precursor - Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: A25011
R:Hitler, I.D., Higgins, C.F.
Eur. J. Biochem. 198, 561-567, 1986
A:Title: Peptide uptake by Salmonella typhimurium. The periplasmic oligopeptide-binding
A:Reference number: A25011; MUID:86274740; PMID:3525163
A:Accession: A25011
A:Molecule type: DNA
A:Residues: 1-542 <Hit>
A:Cross-references: UNIPROT:P06202; UNIPARC:UPI00001703M4; GB:X04194; NID:g47807; PIDN:C
C:Comment: This protein binds peptides up to five amino acids long with high affinity, &
C:Genetics:
A:Gene: oppA
A:Map position: 34 min
C:Superfamily: dipeptide transport protein
C:Keywords: binding protein-dependent transport system; oligopeptide transport; periplas
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-542/Product: oligopeptide-binding protein #status predicted <MAT>

Query Match 40.3%; Score 50; DB 1; Length 542;
Best Local Similarity 58.8%; Pred. No. 6;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 MRNLTKTSLLAGLCTA 17

DB 1 MSNITKSLUAGILTA 17

RESULT 3

T25176
Hypochemical protein T26H5.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25176; T25333
R:Wilkinson, J.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25176
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-378 <Hit>
A:Cross-references: UNIPROT:O62381; UNIPARC:UPI0000061191; EMBL:Z81129; PIDN:CAB03410.1;
A:Experimental source: clone T23F1
R:Gardner, A.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z20017
A:Accession: T25333
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-378 <Hit>
A:Cross-references: UNIPARC:UPI0000061191; EMBL:Z82056; PIDN:CAB04856.1; GSPDB:GN00023;
A:Experimental source: clone T26H5
C:Genetics:
A:Gene: CESP:T26H5.5
A:Map position: 5
A:introns: 42/2; 103/2; 151/2; 311/2

Query Match 39.9%; Score 49.5; DB 2; Length 378;
Best Local Similarity 45.8%; Pred. No. 5.1;
Matches 11; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Oy 1 MRNLTKTSLLAGLCTAQMVFV 23

DB 60 MRLTINAFILGIGICDARMMFI 83

RESULT 4

T41950
glycoprotein H - human herpesvirus 7 (strain J1)

C:Species: human herpesvirus 7

A:Variety: strain J1

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T41950

R:Nicholas, J.

submitted to the EMBL Data Library, December 1995

A:Description: Determination and analysis of the complete nucleotide sequence of human h

A:Reference number: Z22022

A:Accession: T41950

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-690 <NIC>

A:Cross-references: UNIPROT:P52353; UNIPARC:UPI0000138707; EMBL:U43400; PIDN:AAC54710.1

A:Experimental source: strain J1

C:Genetics:

A:Note: U48

C:Superfamily: herpesvirus glycoprotein H

Query Match 39.5%; Score 49; DB 2; Length 690;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Oy 6 KTSLLAGLCTAQMVFVTH 25

DB 438 KIHVLASLGNPLMYFWTH 457

RESULT 5

B84790
Hypochemical protein At2g37230 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84790
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Unayam, L.; Tallon, L.
euss, D.; Mierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84790
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <STO>
A:Cross-references: UNIPROT:Q9ZU03; UNIPARC:UPI000000C710; GB:AE002093; NID:g4056478; PI
C:Genetics:
A:Gene: At2g37230
A:Map position: 2

Query Match 39.5%; Score 49; DB 2; Length 757;
Best Local Similarity 57.9%; Pred. No. 12;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy 3 NLTKTSLLAGLCTAQMVFV 21

DB 359 NATYVSTLPLGLCDAGKVV 377

RESULT 6

A13232
Replication protein A [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: A13232
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ser, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: A13232
A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-405 <KUR>
 A:Cross-references: UNIPROT:066165; UNIPARC:UPI000008EPBA; GB:AE008630; PIDD:AA146279.1;
 A:Experimental source: strain C58 (dupont)
 C:Genetics:
 A:Gene: repA
 A:Genome: plasmid
 C:Superfamily: sopa protein

Query Match 38.7%; Score 48; DB 2; Length 405;
 Best Local Similarity 52.9%; Pred. No. 9.4;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

9 LLLAGLCTAQMVFVTH 25
 |||:|||||:
 266 LTLGLCAATSVIVTH 262

RESULT 7
 T49799
 related to TCM1 protein [imported] - Neurospora crassa

N:Alternate names: protein B1B22.10
 C:Species: Neurospora crassa
 C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
 C:Accession: T49799
 R:Schulte, U.; Altm, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nakamura,
 submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022
 A:Accession: T49799
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-3839 <SCH>
 A:Cross-references: UNIPROT:Q9P4Z1; UNIPARC:UPI000017B515; EMBL:AL356834; GSPDB:GN00116;
 A:Experimental source: BAC clone B1B22; strain OR74A

C:Genetics:
 A:Gene: NCBP:B1B22.10
 A:Map position: 6
 A:Introns: 16/3; 2607/1; 2623/1; 2658/1; 2845/1; 2987/2; 3204/3; 3694/1; 3809/1

Query Match 38.7%; Score 48; DB 2; Length 3839;
 Best Local Similarity 52.6%; Pred. No. 78;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

1 MRRLTKTSLLAGCTPAQ 19
 :|||:|||||:
 679 LRLVNDNSLLGSLCTAR 697

RESULT 8
 AE2089
 phosphoribosylaminoimidazole succinocarboxamide synthase [imported] - Nostoc sp. (strain

C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AE2089

R:Kaneko, T.; Shimpo, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genome Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE2089
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-245 <KUR>
 A:Cross-references: UNIPROT:Q8YUR7; UNIPARC:UPI0000132A86; GB:BA000019; PIDD:BA073967.1;
 A:Experimental source: strain PCC 7120
 C:Genetics:

A:Gene: pncC
 C:Superfamily: phosphoribosylaminoimidazole succinocarboxamide synthase
 Query Match 37.9%; Score 47; DB 2; Length 245;
 Best Local Similarity 50.0%; Pred. No. 8.4;
 Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 4 LTKTSLLAGLCTAQMVFVTH 25
 |||:|||||:
 DB 132 LTRRLTLMELATAEQVEIITH 153

RESULT 9
 T29452
 hypothetical protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T29452
 R:Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1998

A:Reference number: Z20619
 A:Accession: T29452
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-309 <PAR>
 A:Cross-references: UNIPROT:086628; UNIPARC:UPI00000DDAD26; EMBL:AL031155; PIDD:CAA20090
 C:Genetics:
 A:Note: SC3A7.27
 C:Superfamily: Streptomyces coelicolor probable integral membrane protein SC10A7.04

Query Match 37.9%; Score 47; DB 2; Length 309;
 Best Local Similarity 66.7%; Pred. No. 10;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

4 LTKTSLLAGLCTPA 18
 |||:|||||:
 DB 35 LTRITALAGACVNA 49

RESULT 10
 A32812

repA protein - Agrobacterium tumefaciens plasmid pTiB6S3
 C:Species: Agrobacterium tumefaciens
 C>Date: 20-Oct-1989 #sequence_revision 28-Aug-1992 #text_change 09-Jul-2004
 C:Accession: A32812
 R:Tabata, S.; Hooykaas, P.J.J.; Oka, A.
 J. Bacteriol. 171, 1665-1672, 1989

A:Title: Sequence determination and characterization of the replicator region in the tu
 A:Reference number: A32812; MUID:89155477; PMID:2537824

A:Accession: A32812
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-405 <TAB>
 A:Cross-references: UNIPROT:Q52225; UNIPARC:UPI0000084561; GB:M24529; NID:G154803; PIDD
 C:Genetics:
 A:Genome: plasmid
 C:Superfamily: sopa protein

Query Match 37.1%; Score 46; DB 2; Length 405;
 Best Local Similarity 52.9%; Pred. No. 19;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

9 LLLAGLCTAQMVFVTH 25
 |||:|||||:
 DB 266 LTLGLCAATAMVIVTH 282

RESULT 11
 AG0650

periplasmic oligopeptide-binding protein precursor [imported] - Salmonella enterica sub
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A:Note: this species has also been called Salmonella typh
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AG0650
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.W.; Dowd, L.; White, N.; Farrar
 , S.; Moule, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <HEI>
A:Cross-references: UNIPROT:Q9KQV2; UNIPARC:UPI0000063118; GB:AE004265; GB:AE003852; NID:
A:Experimental source: serogroup O1, strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1896
A:Map position: 1

Query Match 36.3%; Score 45; DB 2; Length 472;
Best Local Similarity 76.9%; Pred. No. 32;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MRNLTSTLLAG 13
:|:|:|:|:|:|
16 IRSLTTLSTLLAG 28

RESULT 14
A56382
A:Reference number: A56382; MUID:95238349; PMID:7721763
A:Accession: A56382
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-580 <DIA>
A:Cross-references: UNIPROT:P48777; UNIPARC:UPI0000137940; GB:X79796; NID:g790972; PIDN:
C:Genetics:
A:Gene: uapC
A:Superfamily: uric acid/xanthine transport protein
C:Keywords: transmembrane protein

Query Match 36.3%; Score 45; DB 2; Length 580;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 4 LTKSTLLAGLCTRAQM 20
|:|:|:|:|:|
110 LVSTSLVSGLSAVQM 126

RESULT 15
I64155
A:hypothetical protein HI0638 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: I64155
R:Flatschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glöck, A.; Kelley, J.M.; Weidman, J.;
D.M.; Brandon, R.C.; Pine, L.D.; Fitchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: I64155
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-205 <TRG>
A:Cross-references: UNIPROT:P44796; UNIPARC:UPI0000134674; GB:U2747; GB:U42023; NID:g15
A:Note: Best homolog was a hypothetical protein from Escherichia coli

Query Match 35.5%; Score 44; DB 2; Length 205;
Best Local Similarity 42.9%; Pred. No. 21;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 1 MRNLTSTLLAGLCTRAQM 21

```

DB 1 MKNYHDI VLAAGVCSAKTV 21

Search completed: January 23, 2006, 10:43:07
Job time : 15.6463 secs

The Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:31:19 ; Search time 79.6943 Seconds
(without alignments)
221.323 Million cell updates/sec

Title: US-10-501-838a-12
Perfect score: 124
Sequence: 1 MRNLTKTSLLAGLCTPAQMFVTH 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	324	2	Q6DTN5_9BURK
2	124	100.0	341	2	Q4QZ38_BACSU
3	124	100.0	538	1	NPBB_BACSU
4	52	41.9	207	2	Q8QDD0_METWA
5	51.5	41.5	538	2	Q92P42_RHINE
6	50	40.3	207	2	Q8TRG9_METAC
7	50	40.3	543	1	OPPA_SALTY
8	50	40.3	582	2	Q57NR3_SALCH
9	50	40.3	1003	2	Q7QWK7_GIALA
10	49.5	39.9	378	2	Q62381_GABEL
11	49	39.5	246	2	Q91970_ONCMY
12	49	39.5	246	2	Q91976_ONCMY
13	49	39.5	382	2	Q7NV33_CHRYO
14	49	39.5	483	2	Q8QPK6_9BETA
15	49	39.5	483	2	Q8QPK7_9BETA
16	49	39.5	483	2	Q8QPK8_9BETA
17	49	39.5	483	2	Q8QPK9_9BETA
18	49	39.5	483	2	Q8QPK0_9BETA
19	49	39.5	483	2	Q8QPK1_9BETA
20	49	39.5	543	2	Q65Q82_MANSW
21	49	39.5	690	1	VGLH_HHV7J
22	49	39.5	690	2	Q39950_9BETA
23	49	39.5	690	2	Q56284_9BETA
24	49	39.5	690	2	Q990N1_9BETA
25	49	39.5	690	2	Q990N2_9BETA
26	49	39.5	690	2	Q990N3_9BETA
27	49	39.5	690	2	Q990N4_9BETA
28	49	39.5	690	2	Q990N5_9BETA
29	49	39.5	690	2	Q990N6_9BETA
30	49	39.5	690	2	Q990N7_9BETA
31	49	39.5	690	2	Q990N8_9BETA

32	49	39.5	757	2	Q9ZU13_ARATH	Q9ZU13 arabidopsis
33	48	38.7	297	2	Q4FXU0_LEISHA	Q4FXU0 leishmania
34	48	38.7	404	2	Q6FW26_CANGA	Q6FW26 candida gla
35	48	38.7	405	2	Q66165_9RHIZ	Q66165 agrobacteri
36	48	38.7	405	2	Q8U691_AGRRT5	Q8U691 agrobacteri
37	48	38.7	411	2	Q7D2P0_AGRRT5	Q7D2P0 agrobacteri
38	48	38.7	492	2	Q412E7_GIBZE	Q412E7 gibberella
39	48	38.7	545	2	Q8BHP0_MOUSE	Q8BHP0 mus musculu
40	48	38.7	656	2	Q4HAU0_9BETO	Q4HAU0 deinococcus
41	48	38.7	4065	1	TOM1_NEUCR	Q9P421 neuropept
42	47	37.9	56	2	Q6ACAL_LEIXX	Q6ACAL leifsonia x
43	47	37.9	157	2	Q8GDN5_HEILMO	Q8GDN5 heliobacill
44	47	37.9	203	1	Y1850_PASWU	Q9CJY8 pasteurella
45	47	37.9	245	1	PUR7_AMASP	Q8YUR7 anabaena sp

ALIGNMENTS

RESULT 1
Q6DTN5_9BURK PRELIMINARY; PRT; 324 AA.
ID Q6DTN5_9BURK PRELIMINARY; PRT; 324 AA.
AC Q6DTN5;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Organophosphorus insecticide hydrolase.
GN Name=opdB;
OS Burkholderia sp. FDS-1.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=265945;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FDS-1;
RA Zhang Z., Hong Q., Li S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY646835; AA67170.1; -; Genomic_DNA.
DR SMR; Q6DTN5; 29-322.
DR GO; GO:0016787; F.hydrolase activity; IEA.
DR InterPro; IPR001279; Lactamase-like.
DR Pfam; PF00753; Lactamase_B; 1.
KW Hydrolase.
SQ SEQUENCE 324 AA; 34330 MW; 291D4321A8983AF1 CRC64;
Query Match 100.0%; Score 124; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 MRNLTKTSLLAGLCTPAQMFVTH 25
QY 1 MRNLTKTSLLAGLCTPAQMFVTH 25
Db 1 MRNLTKTSLLAGLCTPAQMFVTH 25
RESULT 2
Q4QZ38_BACSU PRELIMINARY; PRT; 341 AA.
ID Q4QZ38_BACSU PRELIMINARY; PRT; 341 AA.
AC Q4QZ38;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Neutral protease B (Fragment).
GN Name=uprB;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP NOCLEOTIDE SEQUENCE.
RC STRAIN=NGG1;
RA Thang Hong Q., Nguyen Thu H., Nguyen Ngoc Q., Pham Van T.,
RA Duong Thi T.T., Nguyen Tien M., Dinh Duy K.;
RT Cloning and sequencing of the gene encoding neutral protease from
Bacillus subtilis NGG1 isolated from cultivated soil in Vietnam."

```

RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ973636; CAJ01440.1; -; Genomic_DNA.
KM PROTEASE.
FT NON TER
SQ SEQUENCE 341 AA; 38244 MW; ASDBCDB814C98A7 CRC64;
Query Match 100.0%; Score 124; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 1,3e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRNLTSTSLLAGLCTAAQWVFVTH 25
Db 1 MRNLTSTSLLAGLCTAAQWVFVTH 25

RESULT 3
NPRB_BACSU STANDARD; PRT; 538 AA.
AC P39899;
DT 01-FEB-1995 (Rel. 31, last sequence update)
DT 01-FEB-1995 (Rel. 31, last sequence update)
DT 13-SEP-2005 (Rel. 48, last annotation update)
DE Neutral protease B precursor (BC 3.4.24.-).
GN Name=nprb; OrderedLocustNames=BSU11100;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=1423;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 224-228.
RX MEDLINE=92011407; PubMed=1917867;
RA Tran L., Wu X.C., Wong S.L.;
RT "Cloning and expression of a novel protease gene encoding an
RL extracellular neutral protease from Bacillus subtilis.";
RN J. Bacteriol. 173:6364-6372(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=168;
RX MEDLINE=9717785; PubMed=9025291;
RA Levine A., Vannier F., Roche B., Autret S., Mavel D., Serror S.J.;
RT "A 10.3 kbp segment from nprb to argD at the 102 degrees region of the
RL Microbiology 143:175-177(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=168;
RX MEDLINE=9601515; PubMed=9353931;
RA Medina N., Vannier F., Roche B., Autret S., Levine A., Serror S.J.;
RT "Sequencing of regions downstream of addA (98 degrees) and citG (289
RN degrees) in Bacillus subtilis.";
RL Microbiology 143:3305-3308(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=168;
RX MEDLINE=9604403; PubMed=9384377; DOI=10.1038/36786;
RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertoro M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Bourrier L., Brans A., Braum M., Brigelli S.C., Bron S.,
RA Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.-K., Codani J.-J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denliot F., Devine K.M., Diesterhoef A., Ehrlich S.D., Emmerson P.T.,
RA Enlian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.-Y., Glaeser P., Goffeau A., Golightly B.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.-F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klayr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kunano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Manel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogawa B.,
RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,

```

```

RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Sekiguchi J., Sekowska A., Serror S.J., Serror P.,
RA Shin B.-S., Soldo B., Sorokin A., Taccioni B., Takagi T., Takahashi H.,
RA Takemaru K., Takeuchi M., Yamakoshi A., Tanaka T., Terpetra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler H., Wedler H.,
RA Weitznegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasunoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumestein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL subtilis.";
RL Nature 390:249-256(1997).
CC -I- CATALYTIC ACTIVITY: Similar, but not identical, to that of
CC thermolysin.
CC -I- COFACTOR: Binds 4 calcium ions per subunit (Potential).
CC -I- COFACTOR: Binds 1 zinc ion per subunit (Potential).
CC -I- BIOPHYSICOCHEMICAL PROPERTIES:
CC pH dependence:
CC Optimum pH is 6.6;
CC Temperature dependence:
CC Retains 65% activity at 65 degrees Celsius;
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: Belongs to the peptidase M4 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: M62845; AAA22626.1; -; Genomic DNA.
CC EMBL: Z79580; CAB01832.1; -; Genomic DNA.
CC EMBL: Y09476; CA70628.1; -; Genomic DNA.
CC EMBL: Z99109; CAB12950.1; -; Genomic DNA.
CC PIR: A41042; A41042.
CC HSSP: P05806; INPC.
CC MEROPS: M04.012; -.
CC Subtilisin; BGI0691; nprb.
CC InterPro: IPR006025; Pept_M_Zn_BS.
CC InterPro: IPR01570; Peptidase_M4.
CC InterPro: IPR01096; Proppep_M4_M36.
CC InterPro: IPR005075; Proppep_Pepsy.
CC Pfam: PF07504; FTP; 1.
CC Pfam: PF03413; Pepsy; 1.
CC Pfam: PF01447; Peptidase_M4; 1.
CC Pfam: PF02868; Peptidase_M4_C; 1.
CC PRINTS: PR00730; THERMOLYSIN.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC Calcium; Complete proteome; Direct protein sequencing; Hydrolase;
KW Metal-binding; Metalloprotease; Protease; Signal; Zinc; Zymogen.
FT SIGNAL 1 28
FT PROPEP 29 223
FT CHAIN 224 538
FT ACT_SITE 370 370
FT ACT_SITE 453 453
FT METAL 365 365
FT METAL 369 369
FT METAL 373 373
FT METAL 393 393
FT METAL 404 404
FT METAL 406 406
FT METAL 407 407
FT METAL 409 409
FT METAL 412 412
FT METAL 415 415
FT METAL 416 416
FT METAL 419 419
FT METAL 422 422

```

SEQUENCE 538 AA; 59336 MW; 4B7D2B8D8F64F14E CRC64;

Query Match 100.0%; Score 124; DB 1; Length 538;

Best Local Similarity 100.0%; Pred. No. 2e-10; Indels 0; Gaps 0;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MNLTKTSLLAGLCTAAQMVFTVTH 25

1 MNLTKTSLLAGLCTAAQMVFTVTH 25

RESULT 4
Q800D0 METWA PRELIMINARY; PRT; 207 AA.

AC Q800D0; STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;

DT 01-OCT-2002 (TEMBLrel. 22, Created)

DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)

DE Hypothetical protein MM0207.

GN OrderedLocusNames=MM0207;

OS Methanosarcina mazei (Methanosarcina frisia).

OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.

OX NCBI_TaxID=2209;

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;

RX MEDLINE=22120827; PubMed=112125824;

RA Deppehoffer U., Johann A., Hartsch T., Mehl R., Schmitz R.A.,

RA Martiner-Arias R., Henne A., Waeber S., Jacobl C.,

RA Bruggemann H., Leonard T., Christmann A., Boemcke M., Steckel S.,

RA Bhattacharyya A., Lykdis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,

RA Fritz H.-J., Gottschalk G.

RT "The genome of Methanosarcina mazei: evidence for lateral gene

transfer between Bacteria and Archaea."

RL J. Mol. Microbiol. Biotechnol. 4:453-461 (2002).

DR EMBL; AE013244; AAM29903.1; -; Genomic_DNA.

DR InterPro; IPR002716; P1LT_N.

DR InterPro; IPR006596; P1NC.

DR Pfam; PF01850; PIN; 1.

DR SMART; SM00670; P1NC; 1.

KW Complete proteome.

SQ SEQUENCE 207 AA; 24040 MW; 1987D1101B2D2B1 CRC64;

Query Match 41.9%; Score 52; DB 2; Length 207;

Best Local Similarity 52.9%; Pred. No. 12; Indels 0; Gaps 0;

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

9 LLLAGLCTAAQMVFTVTH 25

170 LTFALICRSGEWFVTH 186

RESULT 5
Q92P42 RHIME PRELIMINARY; PRT; 538 AA.

AC Q92P42; STRAIN=1021;

DT 01-DEC-2001 (TEMBLrel. 19, Created)

DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)

DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)

DE PUTATIVE PEPTIDE-BINDING PERIPLASMIC ABC TRANSPORTER PROTEIN.

GN OrderedLocusNames=RO1951; ORFNames=SMC04293;

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium; Ensifer group; Sinorhizobium.

OX NCBI_TaxID=382;

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=1021;

RX MEDLINE=2136507; PubMed=11481430; DOI=10.1073/pnas.161294398;

RA Capela D., Bartley-Hubler F., Gouzy J., Botte G., Ampe F., Batur J.,

RA Boleyard P., Becker A., Boutry M., Cadieu B., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Maury D.,

RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ransperger U.,

RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021."

RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).

DR EMBL; AL591789; CAC46530.1; -; Genomic_DNA.

DR HSPF; P23847; IDP.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; F:transport; IEA.

DR InterPro; IPR000914; SBP_bac_5.

DR Pfam; PF00496; SBP_bac_5; 1.

DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.

KW Complete proteome.

SQ SEQUENCE 538 AA; 59160 MW; 4516D405B34184A0 CRC64;

Query Match 41.5%; Score 51.5; DB 2; Length 538;

Best Local Similarity 65.0%; Pred. No. 33; Indels 5; Gaps 1;

Matches 13; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

1 MNLTKTSLLAGLCTAAQMVFTVTH 20

1 MNLTKTSLLAGLCTAAQMVFTVTH 15

RESULT 6
Q8TKH9 METAC PRELIMINARY; PRT; 207 AA.

AC Q8TKH9; STRAIN=C2A / ATCC 35395 / DSM 2834;

DT 01-JUN-2002 (TEMBLrel. 21, Created)

DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)

DE Predicted protein.

GN OrderedLocusNames=MA3426;

OS Methanosarcina acetivorans.

OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.

OX NCBI_TaxID=2214;

NP NUCLEOTIDE SEQUENCE.

RC STRAIN=C2A / ATCC 35395 / DSM 2834;

RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;

RA Galagan J.E., Nusbaum C., Roy A., Enderizki M.G., MacDonald P.,

RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atencio D., Brown A.,

RA Allen N., Naylor J., Stange-Thomann N., DeArliano K., Johnson R.,

RA Linton L., McEwan P., McEwan K., Talama J., Tirrell A., Ye W.,

RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,

RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,

RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,

RA Springer T.A., Umayam L.A., White O., White R.H., de Maccario E.C.,

RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,

RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,

RA Metcalf W.W., Birren B.

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic

and physiological diversity."

RL Genome Res. 12:532-542 (2002).

DR EMBL; AE011048; AAM06793.1; -; Genomic_DNA.

DR InterPro; IPR002716; P1LT_N.

DR Pfam; PF01850; PIN; 1.

KW Complete proteome.

SQ SEQUENCE 207 AA; 24124 MW; B567AB7B62D7401 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 207;

Best Local Similarity 58.8%; Pred. No. 24; Indels 0; Gaps 0;

Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

9 LLLAGLCTAAQMVFTVTH 25

170 LTFALICRAGLVFTVTH 186

RESULT 7
OPPA_SALTY STANDARD; PRT; 543 AA.

AC P06202; 01-JAN-1988 (Rel. 06, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 13-SEP-2003 (Rel. 48, Last annotation update)

DE Periplasmic oligopeptide-binding protein precursor.

GN Name=oppa; OrderedLocustNames=STM1746;

OS *Salmonella typhimurium*.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; *Salmonella*.

OK NCBI_taxid=602;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=LT2;

RX MEDLINE=88011222; PubMed=2821267;

RT Hiles I.D., Gallagher M.P., Jamieson D.J., Higgins C.F.;

RT "Molecular characterization of the oligopeptide permease of *Salmonella typhimurium*."

RL J. Mol. Biol. 195;125-142(1987).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=86274740; PubMed=3525163;

RA Hiles I.D., Higgins C.F.;

RT "Peptide uptake by *Salmonella typhimurium*. The periplasmic oligopeptide-binding protein."

RL Eur. J. Biochem. 158:561-567(1986).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=LT2 / SGSC1412 / ATCC 700720;

RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali F., Dante M., Du P., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvanev E., Ryan E., Sun H., Flores L., Miller W., Stonking T., Nhan M., Waterston R., Wilson R.K.;

RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2."

RL Nature 413:852-856(2001).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).

RX MEDLINE=94261830; PubMed=8202710;

RA Tame J.R.H., Murshudov G.N., Dodson E.J., Neil T.K., Dodson G.G., Higgins C.F., Wilkinson A.J.;

RT "The structural basis of sequence-independent peptide binding by Oppa protein."

RL Science 264:1578-1581(1994).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).

RX MEDLINE=96363676; PubMed=8747465; DOI=10.1016/S0969-2126(01)00276-3;

RA Tame J.R.H., Dodson E.J., Murshudov G.N., Higgins C.F., Wilkinson A.J.;

RT "The crystal structures of the oligopeptide-binding protein Oppa completed with tripeptide and tetrapeptide ligands."

RL Structure 3:1395-1406(1995).

RN [6]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).

RX MEDLINE=99349693; PubMed=10422831;

RA Davies T.G., Hubbard R.R., Tame J.R.H.;

RT "Relating structure to thermodynamics: the crystal structures and binding affinity of eight Oppa-peptide complexes."

RL Protein Sci. 8:1432-1444(1999).

RN [7]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).

RX MEDLINE=99370085; PubMed=10438628; DOI=10.1006/jmbi.1999.2929;

RA Sleight S.H., Seavers P.R., Wilkinson A.J., Ladbury J.E., Tame J.R.H.;

RT "Crystallographic and calorimetric analysis of peptide binding to Oppa protein."

RL J. Mol. Biol. 291:393-415(1999).

CC -1- FUNCTION: This protein is a component of the oligopeptide permease, a binding protein-dependent transport system, it binds peptides up to five amino acids long with high affinity.

CC -1- SUBCELLULAR LOCATION: Periplasmic.

CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein 5 family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

CC EMBL; X04194; CAA27785.1; -; Genomic DNA.

CC EMBL; X05491; CAA29032.1; -; Genomic DNA.

CC EMBL; AE008777; AAL20664.1; ALT_INIT; Genomic DNA.

CC PIR; A25011; QREBOA.

DR PDB; 1B05; X-ray; A=27-543.

DR PDB; 1B0H; X-ray; A=27-543.

DR PDB; 1B1H; X-ray; A=27-543.

DR PDB; 1B2H; X-ray; A=27-543.

DR PDB; 1B32; X-ray; A=27-543.

DR PDB; 1B3F; X-ray; A=27-543.

DR PDB; 1B3G; X-ray; A=27-543.

DR PDB; 1B3H; X-ray; A=27-543.

DR PDB; 1B3L; X-ray; A/C=27-543.

DR PDB; 1B40; X-ray; A=27-543.

DR PDB; 1B46; X-ray; A=27-543.

DR PDB; 1B4H; X-ray; A=27-543.

DR PDB; 1B4Z; X-ray; A=27-543.

DR PDB; 1B51; X-ray; A=27-543.

DR PDB; 1B52; X-ray; A=27-543.

DR PDB; 1B58; X-ray; A=27-543.

DR PDB; 1B5H; X-ray; A=27-543.

DR PDB; 1B5J; X-ray; A=27-543.

DR PDB; 1B7H; X-ray; A=27-543.

DR PDB; 1B9J; X-ray; A=27-543.

DR PDB; 1JBT; X-ray; A=27-543.

DR PDB; 1JEU; X-ray; A=27-543.

DR PDB; 1JLV; X-ray; A=27-543.

DR PDB; 1JLW; X-ray; A=27-543.

DR PDB; 1OLC; X-ray; A=27-543.

DR PDB; 1OKA; X-ray; A=27-543.

DR PDB; 1OKB; X-ray; A=27-543.

DR PDB; 1RKM; X-ray; A=27-543.

DR PDB; 2OLB; X-ray; A=27-543.

DR PDB; 2RKM; X-ray; A=27-543.

DR StyGene; SG10267; oppa.

DR InterPro; IPR000914; SBP_bac_5.

DR Pfam; PF00496; SBP_bac_5_1.

DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.

KW 3D-structure; Complete proteome; Peptide transport; Periplasmic; Protein transport; Signal; Transport.

FT SIGNAL 1 26

FT DISULFID 27 543

FT CONFLICT 22 24

FT TURN 31 32

FT STRAND 35 35

FT STRAND 40 44

FT TURN 54 55

FT STRAND 59 59

FT HELIX 60 69

FT STRAND 74 76

FT TURN 78 79

FT STRAND 82 84

FT STRAND 87 93

FT TURN 94 96

FT STRAND 97 102

FT TURN 104 105

FT STRAND 107 107

FT TURN 109 110

FT STRAND 113 113

FT HELIX 116 127

FT HELIX 129 131

FT TURN 134 137

FT HELIX 138 141

FT -----

FT Periplasmic oligopeptide-binding protein.

FT SAA -> TP (in Ref. 1 and 2).

```

FT TURN 142 143
PT STRAND 144 144
PT TURN 145 146
PT HELIX 147 151
FT TURN 152 153
FT HELIX 157 159
PT STRAND 162 166
PT TURN 167 168
PT STRAND 169 174
FT TURN 179 180
FT HELIX 181 186
FT HELIX 188 190
FT HELIX 195 201
PT HELIX 202 206
PT TURN 208 210
PT STRAND 217 223
PT TURN 224 226
PT STRAND 227 232
PT TURN 234 235
PT TURN 237 238
PT HELIX 239 241
PT STRAND 246 250
PT HELIX 255 263
PT TURN 264 265
PT STRAND 269 276
PT TURN 276 278
PT HELIX 279 285
PT HELIX 287 289
PT STRAND 290 303
PT TURN 305 306
PT TURN 308 311
PT HELIX 313 322
PT HELIX 325 330
PT TURN 331 333
PT STRAND 338 339
PT TURN 346 347
PT TURN 349 350
PT HELIX 357 360
PT HELIX 375 376
PT TURN 377 378
PT STRAND 386 392
PT HELIX 395 412
PT STRAND 415 421
PT HELIX 423 432
PT TURN 433 433
PT STRAND 437 443
PT TURN 449 449
PT HELIX 450 453
PT HELIX 454 456
PT TURN 458 459
PT TURN 461 462
PT HELIX 470 478
PT TURN 479 481

```

Query Match 40.3%; Score 50; DB 1; Length 543;
 Best Local Similarity 58.8%; Pred. No. 57;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRNLTKTSLIAGLCTA 17
 DB 1 MSNITKSLIAGLITLA 17

RESULT 8
 Q57NR3_BALCH
 ID Q57NR3_BALCH PRELIMINARY; PRT; 582 AA.
 AC Q57NR3;
 DT 10-MAY-2005 (TREMBLERel. 30, Created)
 DT 10-MAY-2005 (TREMBLERel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMBLERel. 30, Last annotation update)
 DE ABC superfamily (Periplasm), oligopeptide transport protein with
 DE chaperone properties.
 GN Name=opa; OrderedLocuNames=SC1742;

OS *Salmonella cholerae-suis* (Salmonella enterica).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI TaxID=591;
 RN (1)
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=SC-B67;
 RX PubMed=15781495;
 RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
 Wang H.-S., Lee Y.-S.;
 RT "The genome sequence of *Salmonella enterica* serovar *Choleraesuis*, a
 highly invasive and resistant zoonotic pathogen.";
 RL Nucleic Acids Res. 33:1690-1698 (2005).
 DR EMBL; AB01720; AX65648.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 582 AA; 65627 MW; E2A4916D7D604850 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 582;
 Best Local Similarity 58.8%; Pred. No. 61;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRNLTKTSLIAGLCTA 17
 DB 40 MSNITKSLIAGLITLA 56

RESULT 9
 Q7QWK7_GIALA
 ID Q7QWK7_GIALA PRELIMINARY; PRT; 1003 AA.
 AC Q7QWK7;
 DT 01-MAR-2004 (TREMBLERel. 26, Created)
 DT 01-MAR-2004 (TREMBLERel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLERel. 26, Last annotation update)
 DS GLP 762 43497 46508.
 OS *Giardia lamblia* ATCC 50803.
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OX NCBI_TaxID=184922;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
 Olsen G.J., Sogin M.L.;
 RT "Draft sequence of the *Giardia lamblia* genome."
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL; AACB0100069; FAJ39455.1; -; Genomic_DNA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004672; F:protein kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00023; Ank_7.
 DR Pfam; PF00069; Kinase; 1.
 DR PRINTS; PRO1415; ANKYRIN.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PSS0297; ANK_REPEAT; 1.
 DR PROSITE; PSS0088; ANK_REPEAT; 2.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 DR ANK repeat; ATP-binding; Kinase; Nucleotide-binding; Repeat;
 KW transference.
 SQ SEQUENCE 1003 AA; 110420 MW; DC824E5D325DF546 CRC64;

Query Match 40.3%; Score 50; DB 2; Length 1003;
 Best Local Similarity 42.1%; Pred. No. 99;
 Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 7 TSLIAGLCTAQMVFVTH 25
 DB 789 TALILAMCGAATVYLH 807

```

RESULT 10
062381 CAEEL PRELIMINARY; PRT; 378 AA.
ID 062381 CAEEL PRELIMINARY; PRT; 378 AA.
AC 062381 CAEEL PRELIMINARY; PRT; 378 AA.
DT 01-NOV-1998 (TRENBLrel. 07, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Hypothetical protein brw-33.
GN Name=brw-33; ORFNames=T26H5.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodietinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology."
RL Science 282:2012-2018 (1998).
CC -1. SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
DR EMBL; Z82056; CAB04856.1; -; Genomic DNA.
DR EMBL; Z81129; CAB03410.1; -; Genomic DNA.
DR EMBL; Z81129; CAB04856.1; JOINED; Genomic DNA.
DR EMBL; Z82056; CAB03410.1; JOINED; Genomic DNA.
DR PIR; T25176; T25176.
DR Ensemble; T26H5.5; Caenorhabditis elegans.
DR WormBase; WBGene0005780; T26H5.5.
DR WormPeP; T26H5.5; CRI8293.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007166; P:G-protein coupled receptor protein signaln. . .; IEA.
DR InterPro; IPR010726; IPR1300.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF06976; DUF1300; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS50262; G_PROTEIN_RECPT_P1_2; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 378 AA; 42911 MW; 6E8E4CA7545896AB CRC64;

Query Match 39.9%; Score 49.5; DB 2; Length 378;
Best Local Similarity 45.8%; Pred. No. 50;
Matches 11; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Cy 1 MENTKTSLLA-GICTAQMVF 23
Db 60 MRTLINAFLGIGICDARWMT 83

RESULT 11
091970 ONCMY PRELIMINARY; PRT; 246 AA.
ID 091970 ONCMY PRELIMINARY; PRT; 246 AA.
AC 091970;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Tumor necrosis factor alpha.
GN Name=tnfa;
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Wang T., Laining K.J., Zou J., Holland J., Hirono I., Aoki T.,
Secombes C.J.;
RL Submitted (May-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ278085; CAB92316.1; -; Genomic DNA.
DR HSRP; P01375; 4TSV.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.

```

```

DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR006053; TNF abc.
DR InterPro; IPR002960; TNF beta.
DR InterPro; IPR006052; TNF family.
DR InterPro; IPR003636; TNF subf.
DR Pfam; PF00229; TNF; 1.
DR PRINTS; PR01234; TNFCROSIPTCT.
DR PRINTS; PR01234; TNFBETA.
DR ProDom; PD002012; TNF subf. 1.
DR SMART; SMO0207; TNF; 1.
DR PROSITE; PS50049; TNF_2; 1.
SQ SEQUENCE 246 AA; 27113 MW; 8CD89DBDF867C6CB CRC64;

Query Match 39.5%; Score 49; DB 2; Length 246;
Best Local Similarity 64.3%; Pred. No. 41;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Cy 9 LLLAGLCTAQMVF 22
Db 41 LLLAGLCAAAALTF 54

RESULT 12
091976 ONCMY PRELIMINARY; PRT; 246 AA.
ID 091976 ONCMY PRELIMINARY; PRT; 246 AA.
AC 091976;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Putative tumor necrosis factor alpha.
GN Name=TNF-alpha;
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21153236; PubMed=11231283;
RA Laining K.J., Wang T., Zou J., Holland J., Hong S., Bole N., Hirono I.,
Aoki T., Secombes C.J.;
RT "Cloning and expression analysis of rainbow trout Oncorhynchus mykiss
tumor necrosis factor-alpha."
RL Eur. J. Biochem. 268:1315-1322 (2001).
DR EMBL; AJ277604; CAB89521.1; -; mRNA.
DR HSRP; P01375; 4TSV.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR006053; TNF abc.
DR InterPro; IPR002960; TNF beta.
DR InterPro; IPR006052; TNF family.
DR InterPro; IPR003636; TNF subf.
DR Pfam; PF00229; TNF; 1.
DR PRINTS; PR01234; TNFCROSIPTCT.
DR PRINTS; PR01234; TNFBETA.
DR ProDom; PD002012; TNF subf. 1.
DR SMART; SMO0207; TNF; 1.
DR PROSITE; PS50049; TNF_2; 1.
FT CHAIN 86 246
SQ SEQUENCE 246 AA; 27113 MW; DCD8971DF867C6C6 CRC64;

Query Match 39.5%; Score 49; DB 2; Length 246;
Best Local Similarity 64.3%; Pred. No. 41;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Cy 9 LLLAGLCTAQMVF 22
Db 41 LLLAGLCAAAALTF 54

RESULT 13
07NV33_CHRVO

```

ID Q7NV33 CHVVO PRELIMINARY; PRT; 382 AA.
 AC Q7NV33;
 DT 01-MAR-2004 (Tremblrel. 26, Created)
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Probable methyl-accepting chemotaxis protein.
 GN OrderedLocusNames=Cv2513;
 OS Chromocacterium violaceum.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 CC Neisseriaceae; Chromobacterium.
 NC NCBITaxid=536;
 OK NCBITaxid=536;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=ATCC 12472 / DSM 30191;
 RX MEDLINE=2282880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
 RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Gilmaraes C.T.,
 RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
 RA Alves-Gomes J.A., Andrade E.M., Araújo J., de Araújo M.F.P.,
 RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
 RA Batista J.S., Belo A., van den Berg C., Bogio M., Bonatto S.,
 RA Bordignon J., Brígido M.M., Brito C.A., Brocchi M., Buritly H.A.,
 RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
 RA Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chuelre L.M.O.,
 RA Cretzynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
 RA Fantiñatti P., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
 RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
 RA Gazzinelli R.T., Gomes E.A., Gonçalves P.R., Grangelro T.B.,
 RA Grattapaglia D., Griard E.C., Hanna B.S., Jardim S.N., Laurino J.,
 RA Leal L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
 RA Madeira H.M.F., Manfio G.P., Maranhão A.Q., Martins W.S.,
 RA di Mauro S.M.Z., de Medeiros S.R.B., Melsener R.V., Moreira M.A.M.,
 RA Nascentimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
 RA Palhao R.F.C., Parente J.A., Pedrosa P.O., Pena S.D.J., Pereira J.O.,
 RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Portich D.P.,
 RA Ramalho-Neco C.B., Reis A.M.M., Rêgo L.U., Rondinelli E.,
 RA Santos E.B.P., Santos F.R., Schneider M.P.C., Senanez H.N., I.C.,
 RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
 RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
 RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmeyni T.,
 RA Vettore A., Wessens R., Zaha A., Simpson A.J.G.;
 RA "The complete genome sequence of Chromobacterium violaceum reveals
 RT remarkable and exploitable bacterial adaptability.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:11650-11655(2003).
 DR EMBL; AEO16918; AA060184.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0006935; P:chemotaxis; IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR004089; ChmTaxis_trained.
 DR Pfam; PF00015; MCPsignal; 1.
 DR PROSITE; PS5011; CHEMOTAXIS_TRANSDUC_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 382 AA; 40450 MW; 7FC17A7542BF614 CRC64;
 Query Match 39.5%; Score 49; DB 2; Length 382;
 Best Local Similarity 46.2%; Pred. No. 60;
 Matches 12; Conservative 7; Mismatches 5; Indels 2; Gaps 1;
 QY 1 MRLUTTSLLAGLCTAQMVFVT 24
 DB 1 MGNRSKSLVFAGLCLASAAALVWVS 26
 RESULT 14
 O80PX6_9BETA PRELIMINARY; PRT; 483 AA.
 ID O80PX6_9BETA PRELIMINARY;
 AC O80PX6;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Glycoprotein H (Fragment).
 OS Human herpesvirus 7.
 OC Human herpesviridae; no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Roseolovirus.
 OK NCBITaxid=10372;
 RN NUCLEOTIDE SEQUENCE.
 RP [1]
 RX MEDLINE=22846754; PubMed=12966550; DOI=10.1002/jmv.10496;
 RA Chan P.K.S., Li C.-K., Chik K.-W., Shing M.M.K., Lee V., Ng K.-C.,
 RA Lam C.-W., Cheung J.L.K., Cheng A.F.;
 RT "Genetic variation of glycoprotein B and H of human herpesvirus 7 in
 RT Hong Kong.";
 RL J. Med. Virol. 71:429-433(2003).
 DR EMBL; AY192549; AA039680.1; -; Genomic_DNA.
 DR InterPro; IPR003493; Herpes_glycopH.
 DR Pfam; PF02489; Herpes_glycop_H; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 483 AA; 56064 MW; 1881E087DE8B8841 CRC64;
 Query Match 39.5%; Score 49; DB 2; Length 483;
 Best Local Similarity 50.0%; Pred. No. 74;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 6 KTSLLAGLCTAQMVFVT 25
 DB 328 KIHVLLASLGNPLEVFWTH 347
 RESULT 15
 O80PX7_9BETA PRELIMINARY; PRT; 483 AA.
 ID O80PX7_9BETA PRELIMINARY;
 AC O80PX7;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Glycoprotein H (Fragment).
 OS Human herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Roseolovirus.
 OK NCBITaxid=10372;
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=22846754; PubMed=12966550; DOI=10.1002/jmv.10496;
 RA Chan P.K.S., Li C.-K., Chik K.-W., Shing M.M.K., Lee V., Ng K.-C.,
 RA Lam C.-W., Cheung J.L.K., Cheng A.F.;
 RT "Genetic variation of glycoprotein B and H of human herpesvirus 7 in
 RT Hong Kong.";
 RL J. Med. Virol. 71:429-433(2003).
 DR EMBL; AY192548; AA039679.1; -; Genomic_DNA.
 DR InterPro; IPR003493; Herpes_glycopH.
 DR Pfam; PF02489; Herpes_glycop_H; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 483 AA; 55992 MW; 620C354FAEBE3CA8 CRC64;
 Query Match 39.5%; Score 49; DB 2; Length 483;
 Best Local Similarity 50.0%; Pred. No. 74;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 6 KTSLLAGLCTAQMVFVT 25
 DB 328 KIHVLLASLGNPLEVFWTH 347
 Search completed: January 23, 2006, 10:40:52
 Job time : 82.6943 secs

The Page Blank (noted)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:52:43 ; Search time 57.4236 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-12
Perfect score: 124
Sequence: 1 MRNLTSTSLLAGLCTAAQWVFVTH 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications, AA, Main:
1: /cgn2_6/pcodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/pcodata/1/pubppaa/US08_PUBCOMB.pep.*
3: /cgn2_6/pcodata/1/pubppaa/US09_PUBCOMB.pep.*
4: /cgn2_6/pcodata/1/pubppaa/US10_PUBCOMB.pep.*
5: /cgn2_6/pcodata/1/pubppaa/US10_PUBCOMB.pep.*
6: /cgn2_6/pcodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	124	100.0	25	4	US-10-665-184-12
2	124	100.0	25	5	US-10-942-300-12
3	124	100.0	25	5	US-10-501-838a-12
4	114	91.9	23	5	US-10-501-838a-72
5	108	87.1	23	4	US-10-665-184-28
6	108	87.1	23	5	US-10-942-300-28
7	108	87.1	23	5	US-10-501-838a-28
8	108	87.1	30	4	US-10-665-184-32
9	108	87.1	30	5	US-10-942-300-32
10	108	87.1	30	5	US-10-501-838a-32
11	50	40.3	272	5	US-10-450-763-33864
12	50	40.3	1059	5	US-10-732-923-12277
13	48	38.7	308	4	US-10-312-088-24
14	48	38.7	308	5	US-10-687-268-24
15	48	38.7	414	4	US-10-343-359-11
16	48	38.7	429	4	US-09-922-501-10
17	48	38.7	429	4	US-10-343-359-10
18	48	38.7	429	5	US-10-787-267a-10
19	47	37.9	203	4	US-10-665-184-60
20	47	37.9	203	5	US-10-942-300-60
21	47	37.9	203	5	US-10-501-838a-54
22	47	37.9	457	4	US-10-312-088-25
23	47	37.9	457	5	US-10-687-268-25
24	46	37.1	109	4	US-10-369-430a-10
25	46	37.1	592	4	US-10-450-763-42557
26	46	37.1	701	3	US-09-978-168-4
27	45	36.3	65	4	US-10-425-115-249144

28	45	36.3	93	4	US-10-767-701-35936	Sequence 35936, A
29	45	36.3	472	4	US-10-156-764-10787	Sequence 10787, A
30	45	36.3	765	6	US-11-097-143-12291	Sequence 15291, A
31	44.5	35.9	348	6	US-11-097-143-126	Sequence 126, App
32	44.5	35.9	506	4	US-10-108-605-265	Sequence 265, App
33	44	35.5	149	4	US-10-437-963-104656	Sequence 104656, A
34	44	35.5	196	4	US-10-437-963-143549	Sequence 143549, A
35	44	35.5	205	4	US-10-665-184-59	Sequence 59, App1
36	44	35.5	205	5	US-10-942-300-59	Sequence 59, App1
37	44	35.5	205	5	US-10-501-838a-53	Sequence 53, App1
38	44	35.5	222	4	US-10-425-115-197731	Sequence 197731, A
39	44	35.5	245	5	US-10-899-557-47	Sequence 47, App1
40	44	35.5	494	4	US-11-097-143-5943	Sequence 5943, App
41	44	35.5	538	4	US-10-156-764-12581	Sequence 12581, A
42	44	35.5	744	3	US-09-833-245-717	Sequence 717, App
43	44	35.5	1146	3	US-09-824-734-2	Sequence 2, App1
44	44	35.5	1146	4	US-10-749-386-2	Sequence 2, App1
45	44	35.5	1146	5	US-10-617-623-9	Sequence 9, App1

ALIGNMENTS

RESULT 1
US-10-665-184-12
; Sequence 12, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-665-184-12
Query Match 100.0%; Score 124; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLTSTSLLAGLCTAAQWVFVTH 25
Db 1 MRNLTSTSLLAGLCTAAQWVFVTH 25

RESULT 2
US-10-942-300-12
; Sequence 12, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 12
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Bacillus subtilis
US-10-942-300-12

Query Match 100.0%; Score 124; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5,1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLTKTSLLAGLCTAAQMVFTVH 25
Db 1 MRNLTKTSLLAGLCTAAQMVFTVH 25

RESULT 3

US-10-501-838A-12
; Sequence 12, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-501-838A-12

Query Match 100.0%; Score 124; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5,1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLTKTSLLAGLCTAAQMVFTVH 25
Db 1 MRNLTKTSLLAGLCTAAQMVFTVH 25

RESULT 4

US-10-501-838A-72
; Sequence 72, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 72
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Bacillus subtilis

US-10-501-838A-72

Query Match 91.9%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 1,7e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NLTKTSLLAGLCTAAQMVFTVH 25
Db 1 NLTKTSLLAGLCTAAQMVFTVH 23

RESULT 5

US-10-665-184-28
; Sequence 28, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-665-184-28

Query Match 87.1%; Score 108; DB 4; Length 23;
Best Local Similarity 95.7%; Pred. No. 1,4e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLTKTSLLAGLCTAAQMVFTV 23
Db 1 MRNLTKTSLLAGLCTAAQMVFTV 23

RESULT 6

US-10-942-300-28
; Sequence 28, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-942-300-28

Query Match 87.1%; Score 108; DB 5; Length 23;
Best Local Similarity 95.7%; Pred. No. 1,4e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;


```
FEATURE:
; OTHER INFORMATION: Synthetic: penetrating peptide
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (27)..(27)
; OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
; OTHER INFORMATION: groups of the lysine residue
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (30)..(30)
; OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
; OTHER INFORMATION: groups of the lysine residue
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (30)..(30)
; OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
; OTHER INFORMATION: peptide via the free amino groups of the lysine residue
US-10-501-838A-32
```

Query Match 87.1%; Score 108; DB 5; Length 30;
Best Local Similarity 95.7%; Pred. No. 1.9e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLKTSLLAGLCTAAQMV 23
Db 1 MRNLKTSLLAGLCTAAQMV 23

```
RESULT 11
US-10-450-763-39684
; Sequence 39684, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hybeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 39684
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (109)..(135)
; OTHER INFORMATION: NO11/NOP2/sun family proteins domain identified by eMATRIX,
; OTHER INFORMATION: accession number BL01153D, p-value=8.322e-14, raw score of 19.69
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (59)..(121)
; OTHER INFORMATION: NO11/NOP2/sun family domain identified by Pfam, accession
; OTHER INFORMATION: name NO11_Nop2_Sun, E-value=0.0015, Pfam score of 15.2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(272)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-39684
```

Query Match 40.3%; Score 50; DB 5; Length 272;
Best Local Similarity 58.8%; Pred. No. 21;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 KTSLLAGLCTAAQMV 21
Db 175 KTSLLAGLCTAAQMV 191

```
RESULT 12
US-10-732-923-12277
; Sequence 12277, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 12277
; LENGTH: 1059
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-10-732-923-12277
```

Query Match 40.3%; Score 50; DB 5; Length 1059;
Best Local Similarity 47.8%; Pred. No. 94;
Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MRNLKTSLLAGLCTAAQMV 23
Db 172 VRNIRHTTLVALLTAFFILV 194

```
RESULT 13
US-10-312-088-24
; Sequence 24, Application US/10312088
; Publication No. US20030219862A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Cogswell, John P.
; APPLICANT: Kabinic, Karen S.
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Martensen, Shelby A.
; APPLICANT: Murdock, Paul R.
; APPLICANT: Smith, Randall F.
; APPLICANT: Strum, Jay C.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Xie, Qing
; APPLICANT: Rizin, Safia K.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029
; CURRENT APPLICATION NUMBER: US/10/312,088
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-312-088-24
```

Query Match 38.7%; Score 48; DB 4; Length 308;
Best Local Similarity 42.1%; Pred. No. 50;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 6 KTSLLAGLCTAAQMV 24
Db 79 EASLIAGGICQASQIFIT 97

RESULT 14

US-10-687-268-24
Sequence 24, Application US/10687268
Publication No. US20050137129A1
GENERAL INFORMATION:
APPLICANT: Agarwal, Pankaj
APPLICANT: Lee, Judithann M.
APPLICANT: Smith, Randall F.
APPLICANT: White, John R.
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GPE0029-1
CURRENT APPLICATION NUMBER: US/10/687,268
CURRENT FILING DATE: 2003-10-15
PRIOR APPLICATION NUMBER: 60/213,161
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 60/213,156
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: PCT/US01/19929
PRIOR FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: 10/312,088
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 24
LENGTH: 308
TYPE: PRT
ORGANISM: Homo sapiens
US-10-687-268-24

Query Match 38.7%; Score 48; DB 5; Length 308;
Best Local Similarity 42.1%; Pred. No. 50;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 6 KTSLLAGLCTAQMVEVT 24

DB 79 EASLLAGLCTAQMVEVT 97

RESULT 15

US-10-343-359-11
Sequence 11, Application US/10343359
Publication No. US2004029234A1
GENERAL INFORMATION:
APPLICANT: Genencor International, Inc.
TITLE OF INVENTION: Enhanced 2-Keto-L-Gluconic Acid
TITLE OF INVENTION: Production
FILE REFERENCE: G6687-PCTA
CURRENT APPLICATION NUMBER: US/10/343,359
CURRENT FILING DATE: 2003-01-20
PRIOR APPLICATION NUMBER: US 09/633,294
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 09/677,032
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 414
TYPE: PRT
ORGANISM: Pantoea citrea
US-10-343-359-11

Query Match 38.7%; Score 48; DB 4; Length 414;
Best Local Similarity 61.1%; Pred. No. 69;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRNLTSLIAGLCTAA 18

DB 64 VRKLVTSLIAGLCTAA 81

Search completed: January 23, 2006, 12:08:58
Job time : 57.4236 secs

This Page Blank (used)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:32:56 ; Search time 13.6463 seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-13
Perfect score: 115
Sequence: 1 IELMIVIAIGILAAIALPAYOEYV 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	111	96.5	154	2	A44809	fimbrial protein -
2	111	96.5	155	2	JI0071	alpha-pilin - Mora
3	111	96.5	156	2	A41490	beta fimbrial prot
4	111	96.5	157	2	A24434	prepilin - Moraxel
5	111	96.5	158	2	A55851	fimbrial protein Q
6	111	96.5	164	2	A42460	pin pilB precurs
7	106	92.2	159	2	S22064	type 4 pilin - Elk
8	106	92.2	159	2	A47699	fimbrial protein c
9	106	92.2	160	2	S40063	fimbrial protein c
10	106	92.2	160	2	S74317	fimbrial protein c
11	106	92.2	160	2	S74318	fimbrial protein c
12	106	92.2	163	2	S40060	fimbrial protein c
13	106	92.2	163	2	S40061	fimbrial protein c
14	106	92.2	163	2	S40062	fimbrial protein c
15	106	92.2	166	1	YONR6	fimbrial protein c
16	106	92.2	166	1	S55496	fimbrial protein c
17	106	92.2	167	1	S03091	fimbrial protein c
18	106	92.2	167	1	A29611	fimbrial protein p
19	106	92.2	168	2	S57418	fimbrial protein M
20	106	92.2	168	2	S22102	fimbrial protein M
21	106	92.2	168	2	S57416	fimbrial protein M
22	106	92.2	168	2	S22103	fimbrial protein M
23	106	92.2	169	2	S22104	fimbrial protein M
24	106	92.2	170	2	D82021	fimbrial protein p
25	106	92.2	170	2	F81246	pin pilB NMB0018
26	106	92.2	180	2	S15327	fimbrial protein M
27	106	92.2	214	2	S15326	fimbrial protein M
28	105	91.3	157	2	A31105	fimbrial protein p
29	102	88.7	149	2	A25023	type 4 fimbrial pr

30	102	88.7	150	1	YOBSPA
31	102	88.7	150	2	B24603
32	102	88.7	150	2	A43504
33	102	88.7	156	2	S15638
34	102	88.7	157	2	S15637
35	102	88.7	157	2	S15642
36	102	88.7	159	2	S15640
37	102	88.7	159	2	S15639
38	102	88.7	164	2	S15641
39	102	88.7	168	2	S22107
40	101	87.8	85	2	B24886
41	101	87.8	136	2	S52692
42	101	87.8	156	2	S15267
43	101	87.8	156	2	S15266
44	101	87.8	157	2	A46566
45	101	87.8	160	2	A37167

ALIGNMENTS

RESULT 1
A44809
fimbrial protein - Moraxella nonliquefaciens
N/Alternate names: 198 antigen; pilin
C/Species: Moraxella nonliquefaciens
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: A44809; A05122
R/Tonjum, T.; Marrs, C.F.; Rozsa, F.; Bovre, K.
J. Gen. Microbiol. 137, 2483-2490, 1991
A/Title: The type 4 pilin of Moraxella nonliquefaciens exhibits unique similarities with
A/Reference number: A44809; MUID:92121902; PMID:1770363
A/Accession: A44809
A/Molecule type: DNA
A/Residues: 1-154 <TON>
A/Cross-references: UNIPROT:P09829; UNIPARC:UP1000012AA7
A/Experimental source: strain NCCTC 7784
A/Note: sequence extracted from NCBI backbone (NCBIN:77471, NCBIPI:77472)
R/Proholm, L.O.; Sletten, K.
FEBS Lett. 73, 29-32, 1977
A/Title: Purification and N-terminal sequence of a fimbrial protein from Moraxella nonl.
A/Reference number: A05122; MUID:77116125; PMID:838045
A/Accession: A05122
A/Molecule type: protein
A/Residues: 7-55 <FRO>
A/Cross-references: UNIPARC:UP100001781P3
C/Genetics:
A/Gene: tfrp
C/Superfamily: gonococcal fimbrial protein
C/Keywords: methylated amino end; surface antigen
F/7/Modified site: methylated amino end (Phe) (in mature form) #status experimental

Query Match 96.5%; Score 111; DB 2; Length 154;
Best Local Similarity 92.0%; Pred. No. 7.4e-08;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVIAIGILAAIALPAYOEYV 25
DB 10 IELMIVIAIGILAAIALPAYODYI 34

RESULT 2
JI0071
alpha-pilin - Moraxella bovis (strain Bpp63)
C/Species: Moraxella bovis
A/Note: host Bos primigenius taurus (cattle)
C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C/Accession: JI0071
R/Rueh, W.W.; Marrs, C.F.; Fernandez, R.; Faltow, S.; Schoolnik, G.K.
J. Exp. Med. 168, 983-1002, 1988
A/Title: Purification, characterization, and pathogenicity of Moraxella bovis pil.
A/Reference number: JI0071; MUID:89010522; PMID:2802184
A/Accession: JI0071

A:Molecule type: protein
A:Residues: 1-155 <RUE>
A:Crosss-references: UNIPROT:P20657, UNIPARC:UPI00001781F5
A:Note: about 50% of the amino-terminus is N-methylated; the remaining 50% of the molecule is N-methylated
A:Note: unpublished DNA sequence evidence indicated 117-Thr and 120-Thr, which were not C:Keywords: methylated amino end
C:Superfamily: gonococcal fimbrial protein
P:11-155/Product: alpha-pilin 1 #status experimental <MAT1>
P:12-155/Product: alpha-pilin 2 #status experimental <MAT2>
P:11/Modified site: methylated amino end (Phe) #status experimental
P:131-150/Disulfide bonds: #status experimental

Query Match 96.5%; Score 111; DB 2; Length 155;
Best Local Similarity 92.0%; Pred. No. 7.5e-08;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIIGILAAIALPAYQEVY 25
Db 4 IELMIVIAIIGILAAIALPAYQDYI 28
|||||
|||||

RESULT 3
A:1490
pilin precursor - Moraxella bovis
C:Species: Moraxella bovis
C:Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C:Accession: A14190
R:Elleman, T.C.; Hoynes, P.A.; Leggett, A.W.D.
Infect. Immun. 58, 1678-1684, 1990
A:Title: Characterization of the pilin gene of Moraxella bovis Dalton 2d and expression
A:Reference number: A14190; MUID:90256243; PMID:1971258
A:Accession: A14190
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <ELL>
A:Crosss-references: UNIPROT:Q59507, UNIPARC:UPI00000B2378, GB:M92155; GB:M2665; NID:g149760; PIDN:
C:Superfamily: gonococcal fimbrial protein
C:Keywords: methylated amino end
P:7/Modified site: methylated amino end (Phe) (in mature form) #status predicted

Query Match 96.5%; Score 111; DB 2; Length 156;
Best Local Similarity 92.0%; Pred. No. 7.5e-08;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIIGILAAIALPAYQEVY 25
Db 10 IELMIVIAIIGILAAIALPAYQDYI 34
|||||
|||||

RESULT 4
A:24434
beta fimbrial protein precursor - Moraxella bovis
A:Alternate names: beta pilin
C:Species: Moraxella bovis
C:Date: 08-Aug-1987 #sequence_revision 08-Aug-1987 #text_change 09-Jul-2004
C:Accession: A24434; J10072
R:Maris, C.F.; Schoonink, G.; Koomey, J.M.; Hardy, J.; Rothbard, J.; Falkow, S.
J. Bacteriol. 163, 132-139, 1985
A:Reference number: A24434; MUID:85234350; PMID:2861194
A:Accession: A24434
A:Molecule type: DNA
A:Residues: 1-157 <MAR>
A:Crosss-references: UNIPROT:P07640, UNIPARC:UPI000012AAC7, GB:M11435; NID:g149760; PIDN:
A:Experimental source: strain EPP63
R:Ruehl, W.W.; Maris, C.F.; Fernandez, R.; Falkow, S.; Schoonink, G.K.
J. Exp. Med. 168, 983-1002, 1988
A:Title: Purification, characterization, and pathogenicity of Moraxella bovis pilin.
A:Reference number: J10071; MUID:89010522; PMID:2902184
A:Accession: J10072
A:Molecule type: protein
A:Residues: 7-86,93-122 <RUE>
A:Crosss-references: UNIPARC:UPI00001781F2
A:Note: about 50% of the amino-terminus is N-methylated; the remaining 50% of the molecule is N-methylated

```

C:Superfamily: gonococcocal fimbrial protein
C:Keywords: methylated amino end; surface antigen
F:1-6/Domain: propeptide #status predicted <PRO>
F:7-157/Product: beta fimbrial protein 1 #status experimental <MNT>
F:8-157/Product: beta fimbrial protein 2 #status experimental <MA2>
F:7/Modified site: methylated amino end (Phe) (in mature form) #status experimental
F:136-155/Disulfide bonds: #status experimental

Query Match          96.5%; Score 111, DB 2; Length 157;
Best Local Similarity 92.0%; Pred. No. 7.5e-08;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IELMTVIAITIGILAAIALPAYQERY 25
         |||||
DB       10 IELMTVIAITIGILAAIALPAYQDYI 34

RESULT 5
A55851
Prepilin - Moraxella bovis
C:Species: Moraxella bovis
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: A55851
R:Atwell, J.L.; Tennent, J.M.; Lepper, A.W.; Eklleman, T.C.
J. Bacteriol. 176, 4875-4882, 1994
A>Title: Characterization of pilin genes from seven serologically defined prototype strai
A:Reference number: A55851; NCID:94327452; PMID:8051000
A:Accession: A55851
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-158 <RES>
A:Cross-references: UNIPROT:Q59501; UNIPARC:UP100000BACSD; GB:L32965; NID:g488303; PIDN:
C:Superfamily: gonococcocal fimbrial protein
C:Keywords: methylated amino end
F:7/Modified site: methylated amino end (Phe) (in mature form) #status predicted

Query Match          96.5%; Score 111, DB 2; Length 158;
Best Local Similarity 92.0%; Pred. No. 7.6e-08;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IELMTVIAITIGILAAIALPAYQERY 25
         |||||
DB       10 IELMTVIAITIGILAAIALPAYQDYI 34

RESULT 6
A42460
fimbrial protein O - Moraxella lacunata (ATCC 17956) plasmid pmxLJ
C:Species: Moraxella lacunata
C:Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 29-Jan-1999
C:Accession: A42460
R:Rozaa, F.W.; Marrs, C.F.
J. Bacteriol. 173, 4000-4006, 1991
A>Title: Interrelating sequence differences between the pilin gene inversion regions of Mo
A:Reference number: A42460; NCID:91286182; PMID:2061282
A:Accession: A42460
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-164 <ROZ>
A:Cross-references: UNIPARC:UP100001781F4; GB:M59711
C:Genetics:
A:Genome: plasmid
C:Superfamily: gonococcocal fimbrial protein
C:Keywords: methylated amino end
F:7/Modified site: methylated amino end (Phe) (in mature form) #status predicted

Query Match          96.5%; Score 111, DB 2; Length 164;
Best Local Similarity 92.0%; Pred. No. 7.8e-08;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IELMTVIAITIGILAAIALPAYQERY 25
         |||||
DB       10 IELMTVIAITIGILAAIALPAYQDYI 34

```


RESULT 7

822064
piliin piliE precursor - Neisseria gonorrhoeae (fragment)
C/Species: Neisseria gonorrhoeae
A/Variety: strain MS11
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: S22064
R/Haas, R.
Submitted to the EMBL Data Library, June 1992
A/Reference number: S22064
A/Accession: S22064
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-159 <HAA>
A/Cross-references: UNIPROT:Q51321; UNIPARC:UPI00000BD9D7; EMBL:X66620; NID:945254; PIDD
C/Superfamily: gonococcal fimbrial protein

Query Match

Best Local Similarity 92.2%; Score 106; DB 2; Length 159;
Best Local Similarity 87.5%; Pred. No. 3.6e-07;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY

1 IELMIVAIIGILAAALPAYQDY 24
|||||:||||:||||:||||:|
4 IELMIVAIIGILAAALPAYQDY 27

RESULT 8

A47699
type 4 piliin - Bikenella corrodens
N/Alternate names: N-methylphenylalanine-type piliin
C/Species: Bikenella corrodens
C/Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A47699; S23844
R/Rao, V.K.; Progleke-Fox, A.
J. Gen. Microbiol. 139, 651-660, 1993
A/Title: Cloning and sequencing of two type 4 (N-methylphenylalanine) piliin genes from B.
A/Reference number: A47699; PMID:93232782; PMID:8473871
A/Accession: A47699
A/Molecule type: DNA
A/Residues: 1-159 <RAO>
A/Cross-references: UNIPROT:P35645; UNIPARC:UPI0000129B89; EMBL:Z12609; NID:941329; PIDD
A/Experimental source: ATCC 23834
A/Note: sequence extracted from NCBI backbone (NCBI:P.130001)

C/Genetics:
A/Gene: ecpa
C/Superfamily: gonococcal fimbrial protein
C/Keywords: methylated amino end

F/Modified site: methylated amino end (Phe) (in mature form) #status predicted

Query Match

Best Local Similarity 92.2%; Score 106; DB 2; Length 159;
Best Local Similarity 88.0%; Pred. No. 3.6e-07;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY

1 IELMIVAIIGILAAALPAYQDY 25
|||||:||||:||||:||||:|
11 IELMIVAIIGILAAALPAYQDY 35

RESULT 9

S40063
fimbrial protein class I - Neisseria meningitidis (fragment)
N/Alternate names: piliin
C/Species: Neisseria meningitidis
A/Variety: strain C311 variant 28; strain C311 variant 3
C/Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S40063; S74316
R/Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A/Title: Pilius-facilitated adherence of Neisseria meningitidis to human epithelial and
ence and the glycosylation status of piliin.
A/Reference number: S40060; PMID:95020562; PMID:7934852

A/Accession: S40063
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-160 <VIR>

A/Cross-references: UNIPROT:Q57135; UNIPARC:UPI00000BA7AA; EMBL:L22678; NID:9349523; PII
A/Experimental source: strain C311 variant 28

A/Accession: S74316

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA
A/Residues: 1-160 <VIR>

A/Cross-references: UNIPARC:UPI00000BA7AA; EMBL:L22636; NID:9349013; PIDD:AAA67334.1; PII
A/Experimental source: strain C311 variant 3

C/Genetics:
A/Gene: piliE
C/Superfamily: gonococcal fimbrial protein
C/Keywords: fimbria

Query Match

Best Local Similarity 92.2%; Score 106; DB 2; Length 160;
Best Local Similarity 87.5%; Pred. No. 3.6e-07;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY

1 IELMIVAIIGILAAALPAYQDY 24
|||||:||||:||||:||||:|
4 IELMIVAIIGILAAALPAYQDY 27

RESULT 10

S74317
fimbrial protein class I - Neisseria meningitidis (strain C311 variant 7) (fragment)
N/Alternate names: piliin
C/Species: Neisseria meningitidis
A/Variety: strain C311 variant 7
C/Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S74317
R/Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A/Title: Pilius-facilitated adherence of Neisseria meningitidis to human epithelial and
ence and the glycosylation status of piliin.
A/Reference number: S40060; PMID:95020562; PMID:7934852
A/Accession: S74317
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-160 <VIR>
A/Cross-references: UNIPROT:Q51234; UNIPARC:UPI00000AFDPI; EMBL:L22676; NID:9349519; PII

C/Genetics:
A/Gene: piliE
C/Superfamily: gonococcal fimbrial protein
C/Keywords: fimbria

Query Match

Best Local Similarity 92.2%; Score 106; DB 2; Length 160;
Best Local Similarity 87.5%; Pred. No. 3.6e-07;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY

1 IELMIVAIIGILAAALPAYQDY 24
|||||:||||:||||:||||:|
4 IELMIVAIIGILAAALPAYQDY 27

RESULT 11

S74318
fimbrial protein class I - Neisseria meningitidis (strain C311 variant 16) (fragment)
N/Alternate names: piliin
C/Species: Neisseria meningitidis
A/Variety: strain C311 variant 16
C/Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S74318
R/Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A/Title: Pilius-facilitated adherence of Neisseria meningitidis to human epithelial and
ence and the glycosylation status of piliin.
A/Reference number: S40060; PMID:95020562; PMID:7934852
A/Accession: S74318
A/Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 1-160 <VIR>
A:Cross-references: UNIPROT:Q51235, UNIPARC:UPI000008B820, EMBL:L22677, NID:g349521, PID
C:Genetics:
A:Gene: p11E
C:Superfamily: gonococcal fimbrial protein
C:Keywords: fimbria

Query Match 92.2%; Score 106; DB 2; Length 160;
Best Local Similarity 87.5%; Pred. No. 3.6e-07;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGTILAAVLPAYQDY 24
Db 4 IELMIVIAIGTILAAVLPAYQDY 27

RESULT 12
S40060
fimbrial protein class I - Neisseria meningitidis (strain MCS8 variant 5) (fragment)
N:Alternate names: p11n
C:Species: Neisseria meningitidis
A:Variety: strain MCS8 variant 5
C:Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S40060
R:Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A:Title: Pilus-facilitated adherence of Neisseria meningitidis to human epithelial and e
ence and the glycosylation status of p11n.
A:Reference number: S40060; MUID:95020562; PMID:7934852
A:Accession: S40060
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-163 <VIR>
A:Cross-references: UNIPROT:Q51237, UNIPARC:UPI000008B823, EMBL:L22681, NID:g349015, PID
C:Genetics:
A:Gene: p11E
C:Superfamily: gonococcal fimbrial protein
C:Keywords: fimbria

Query Match 92.2%; Score 106; DB 2; Length 163;
Best Local Similarity 87.5%; Pred. No. 3.6e-07;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGTILAAVLPAYQDY 24
Db 4 IELMIVIAIGTILAAVLPAYQDY 27

RESULT 13
S40061
fimbrial protein class I - Neisseria meningitidis (strain MCS8 variant 6) (fragment)
N:Alternate names: p11n
C:Species: Neisseria meningitidis
A:Variety: strain MCS8 variant 6
C:Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
R:Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A:Title: Pilus-facilitated adherence of Neisseria meningitidis to human epithelial and e
ence and the glycosylation status of p11n.
A:Reference number: S40060; MUID:95020562; PMID:7934852
A:Accession: S40061
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-163 <VIR>
A:Cross-references: UNIPROT:Q51236, UNIPARC:UPI000008F88D, EMBL:L22679, NID:g349525; PID
C:Genetics:
A:Gene: p11E
C:Superfamily: gonococcal fimbrial protein
C:Keywords: fimbria

Query Match 92.2%; Score 106; DB 2; Length 163;

Best Local Similarity 87.5%; Pred. No. 3.6e-07;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGTILAAVLPAYQDY 24
Db 4 IELMIVIAIGTILAAVLPAYQDY 27

RESULT 14
S40062
fimbrial protein class I - Neisseria meningitidis (strain MCS8 variant 7) (fragment)
N:Alternate names: p11n
C:Species: Neisseria meningitidis
A:Variety: strain MCS8 variant 7
C:Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S40062
R:Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A:Title: Pilus-facilitated adherence of Neisseria meningitidis to human epithelial and e
ence and the glycosylation status of p11n.
A:Reference number: S40060; MUID:95020562; PMID:7934852
A:Accession: S40062
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-163 <VIR>
A:Cross-references: UNIPROT:Q57441, UNIPARC:UPI000008B81C, EMBL:L22683, NID:g349527; PID
C:Genetics:
A:Gene: p11E
C:Superfamily: gonococcal fimbrial protein
C:Keywords: fimbria

Query Match 92.2%; Score 106; DB 2; Length 163;
Best Local Similarity 87.5%; Pred. No. 3.6e-07;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGTILAAVLPAYQDY 24
Db 4 IELMIVIAIGTILAAVLPAYQDY 27

RESULT 15
Y0NMG
fimbrial protein class I precursor - Neisseria gonorrhoeae
N:Alternate names: MS11 antigen; p11n
C:Species: Neisseria gonorrhoeae
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004
C:Accession: A94007; A92776; A44149; S13372; C24886; A24886; E24886; A03498
R:Meyer, T.F.; Billiard, E.; Haas, R.; Scorzbach, S.; So, M.
Proc. Natl. Acad. Sci. U.S.A. 81, 6110-6114, 1984
A:Title: Pilus genes of Neisseria gonorrhoeae: chromosomal organization and DNA sequence.
A:Reference number: A94007; MUID:85014915; PMID:6148752
A:Accession: A94007
A:Molecule type: DNA
A:Residues: 1-166 <MEX>
A:Cross-references: UNIPROT:P02974, UNIPARC:UPI000016F8F7, GB:K02078, NID:g150284, PIDN:
A:Experimental source: strain MS11
R:Schoolnik, G.K.; Fernandez, R.; Tai, J.Y.; Rothbard, J.; Gotschlich, E.C.
J. Exp. Med. 159, 1351-1370, 1984
A:Title: Gonococcal p11. Primary structure and receptor binding domain.
A:Reference number: A92776; MUID:84187404; PMID:6143785
A:Accession: A92776
A:Molecule type: protein
A:Residues: 8-166 <SCH>
A:Cross-references: UNIPARC:UPI0000141D34
A:Experimental source: strain MS11
A:Note: phosphoserine was detected but not located
R:Johnson, A.B.; Pfeiffer, J.; Normark, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 3204-3208, 1992
A:Title: Neisseria gonorrhoeae p11c expression provides a selective mechanism for struct
A:Reference number: A44149; MUID:92228752; PMID:11348857
A:Accession: A44149
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA

A/Residues: 1-166 <JON1>
 A/Cross-references: UNIPARC:UPI000016FBFP
 A/Experimental source: strain MS1mk [P+]-u
 R/Jonsson, A.B.; Nyberg, G.; Normark, S.
 EMBO J. 10, 477-488, 1991
 A/Title: Phase variation of gonococcal pil1 by frameshift mutation in pilC, a novel gene
 A/Reference number: S13372; MUID:9112056; PMID:1671354
 A/Accession: S13372
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-166 <JON2>
 A/Cross-references: UNIPARC:UPI000016FBFP
 R/Bergstrom, S.; Robbins, K.; Komey, J.M.; Swanson, J.
 Proc. Natl. Acad. Sci. U.S.A. 83, 3890-3894, 1986
 A/Title: Pilaction control mechanisms in *Neisseria gonorrhoeae*.
 A/Reference number: A24886; MUID:8623338; PMID:2872674
 A/Accession: C24886
 A/Molecule type: DNA
 A/Residues: 1-24, 'L', 26-166 <BER1>
 A/Cross-references: UNIPARC:UPI00001747DF; GB:M13222; NID:G150290
 A/Experimental source: P(-)rp(-) revertant P(+) rev
 A/Accession: A24886
 A/Molecule type: DNA
 A/Residues: 1-24, 'L', 26-60, 'N', 62, 'T', 64-66, 'A', 70-73, 'G', 75, 'S', 77, 'T', 79-158, 'L', 160-1
 A/Cross-references: UNIPARC:UPI00001747E0; GB:M13222; NID:G150290
 A/Experimental source: P(++)
 A/Accession: E24886
 A/Molecule type: DNA
 A/Residues: 1-24, 'L', 26-60, 'N', 62, 'T', 64-66, 'A', 70-85, 'QK', 88-90, 'AK', 93-99, 'A', 101-102,
 A/Cross-references: UNIPARC:UPI00001747B1; GB:M13222; NID:G150290
 A/Experimental source: P(-)rp(+) revertant P(++)
 C/Function:
 A/Description: surface protein that binds various host cell molecules
 C/Superfamily: gonococcal fimbrial protein
 C/Keywords: fimbria; glycoprotein; methylated amino end; phosphoprotein; surface antigen
 P/1-7/Domain: propeptide #status predicted <PRO>
 P/8-166/Product: fimbrial protein #status predicted <MAT>
 P/31-111/Domain: receptor binding #status predicted <RBC>
 P/8/Modified site: methylated amino end (Phe) (in mature form) #status experimental
 P/101/Binding site: sn-1-glycerolphosphate (Ser) (covalent) #status predicted
 P/128-158/Disulfide bonds: #status experimental

Query Match 92.2%; Score 106; DB 1; Length 166;
 Best local similarity 87.5%; Pred. No. 3.7e-07;
 Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVIAITIGIIAIALPAYQRY 24
 |||||:|||||:|||||:
 DB 11 IELMIVIAIVGIIAVALPAYQDY 34

Search completed: January 23, 2006, 10:43:07
 Job time : 13.6463 secs

The page edge cut

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:35:50 ; Search time 22.9258 seconds
(without alignments)
90.156 Million cell updates/sec

Title: US-10-501-838a-13
Perfect score: 115
Sequence: 1 IELMIVAIIGILAIAPAYOEYV 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/6.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/H.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/RTUS.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/RC.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfillset.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	93.0	169	US-09-328-352-6277	Sequence 6277, App
2	106	92.2	166	US-09-350-841A-1595	Sequence 1595, App
3	106	92.2	167	US-08-486-099-109	Sequence 109, App
4	106	92.2	167	US-08-360-107A-119	Sequence 119, App
5	106	92.2	167	US-08-484-233B-109	Sequence 109, App
6	106	92.2	167	US-08-475-668A-109	Sequence 109, App
7	106	92.2	167	US-08-475-668A-109	Sequence 109, App
8	106	92.2	167	US-08-485-513A-109	Sequence 109, App
9	106	92.2	167	US-08-471-913A-109	Sequence 109, App
10	106	92.2	167	US-08-485-264A-109	Sequence 109, App
11	106	92.2	167	US-08-474-349A-109	Sequence 109, App
12	106	92.2	167	US-08-470-896-109	Sequence 109, App
13	106	92.2	167	US-08-485-546A-109	Sequence 109, App
14	106	92.2	167	US-08-487-266A-109	Sequence 109, App
15	106	92.2	167	US-08-484-741-109	Sequence 109, App
16	106	92.2	167	US-09-252-991A-26702	Sequence 26702, A
17	101	87.8	26	US-08-891-254-9	Sequence 9, App1
18	101	87.8	26	US-08-819-539-9	Sequence 9, App1
19	101	87.8	26	US-09-030-270A-9	Sequence 9, App1
20	101	87.8	26	US-08-984-207-9	Sequence 9, App1
21	101	87.8	26	US-09-013-587-9	Sequence 9, App1
22	101	87.8	26	US-09-086-118-29	Sequence 29, App1
23	101	87.8	26	US-09-431-614-17	Sequence 17, App1
24	101	87.8	26	US-09-412-100-37	Sequence 37, App1
25	101	87.8	26	PCT-US96-08819-9	Sequence 9, App1
26	100	87.0	154	US-08-486-099-108	Sequence 108, App
27	100	87.0	154	US-08-360-107A-118	Sequence 118, App

28	100	87.0	154	2	US-08-484-223B-108	Sequence 108, App
29	100	87.0	154	2	US-08-919-597-108	Sequence 108, App
30	100	87.0	154	2	US-08-475-668A-108	Sequence 108, App
31	100	87.0	154	2	US-08-485-551A-108	Sequence 108, App
32	100	87.0	154	2	US-08-471-913A-108	Sequence 108, App
33	100	87.0	154	2	US-08-485-264A-108	Sequence 108, App
34	100	87.0	154	2	US-08-474-349A-108	Sequence 108, App
35	100	87.0	154	2	US-08-470-896-108	Sequence 108, App
36	100	87.0	154	2	US-08-485-546A-108	Sequence 108, App
37	100	87.0	154	2	US-08-487-266A-108	Sequence 108, App
38	100	87.0	154	2	US-08-484-741-108	Sequence 108, App
39	90	78.3	162	2	US-09-303-518D-340	Sequence 340, App
40	87	75.7	178	2	US-09-489-039A-12339	Sequence 12339, A
41	87	75.7	292	2	US-09-199-637A-419	Sequence 419, App
42	86	74.8	162	2	US-09-303-518D-344	Sequence 344, App
43	85	73.9	256	2	US-09-252-991A-21475	Sequence 21475, A
44	84	73.0	163	2	US-09-543-681A-4903	Sequence 4903, App
45	83	72.2	144	2	US-09-328-352-6283	Sequence 6283, App

ALIGNMENTS

RESULT 1
US-09-328-352-6277
Sequence 6277, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6277
LENGTH: 169
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6277

Query Match 93.0%; Score 107; DB 2; Length 169;
Best Local Similarity 88.0%; Pred. No. 7.9e-09;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IELMIVAIIGILAIAPAYOEYV 25
DB 26 IELMIVAIIGILAIAPAYONYV 50

RESULT 2
US-09-350-841A-1595
Sequence 1595, Application US/09350841A
Patent No. 6750008
GENERAL INFORMATION:
APPLICANT: Jette, Peter;
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
FILE REFERENCE: 7872-066-999
CURRENT FILING DATE: 1999-07-09
NUMBER OF SEQ ID NOS: 1946
SOFTWARE: PatentIn-Ver 2.1
SEQ ID NO 1595
LENGTH: 166
TYPE: PRT
ORGANISM: Human immunodeficiency virus type 1
US-09-350-841A-1595

Query Match 92.2%; Score 106; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TELMTVIAIGTILAAALPAYOEY 24
|||||:|||||:|||||:|||||:
Db 11 TELMTVIAIGTILAAALPAYODY 34

RESULT 3
US-08-486-099-109
; Sequence 109, Application US/08486099
; Patent No. 6013263
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; TITLE OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,099
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-486-099-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TELMTVIAIGTILAAALPAYOEY 24
|||||:|||||:|||||:|||||:
Db 11 TELMTVIAIGTILAAALPAYODY 34

RESULT 4
US-08-360-107A-119
; Sequence 119, Application US/08360107A
; Patent No. 6017536
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.

APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

US-08-484-223B-109
; Sequence 109, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

Oy 1 TELMTVIAIGTILAAALPAYOEY 24
|||||:|||||:|||||:|||||:
Db 11 TELMTVIAIGTILAAALPAYODY 34

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TELMTVIAIGTILAAALPAYOEY 24
|||||:|||||:|||||:|||||:
Db 11 TELMTVIAIGTILAAALPAYODY 34

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,223B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-223B-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIIGILAAALPAYOEY 24
DB 11 IELMIVAIIGILAAALPAYODY 34

RESULT 6
US-08-919-597-109
Sequence 109, Application US/08919597
Patent No. 6054265
GENERAL INFORMATION:
APPLICANT: Bolognesi, Daniel P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-919-597-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIIGILAAALPAYOEY 24
DB 11 IELMIVAIIGILAAALPAYODY 34

RESULT 7
US-08-475-668A-109
Sequence 109, Application US/08475668A
Patent No. 6060065
GENERAL INFORMATION:
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-475-668A-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIIGILAAALPAYOEY 24
DB 11 IELMIVAIIGILAAALPAYODY 34

TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-919-597-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIIGILAAALPAYOEY 24
DB 11 IELMIVAIIGILAAALPAYODY 34

RESULT 7
US-08-475-668A-109
Sequence 109, Application US/08475668A
Patent No. 6060065
GENERAL INFORMATION:
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-475-668A-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIIGILAAALPAYOEY 24
DB 11 IELMIVAIIGILAAALPAYODY 34

RESULT 8
US-08-485-551A-109
; Sequence 109, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,551A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-551A-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGILAAALPAYOEY 24
Db 11 IELMIVIAIGILAAALPAYODY 34

RESULT 9
US-08-471-913A-109
; Sequence 109, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE

; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,913A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-471-913A-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGILAAALPAYOEY 24
Db 11 IELMIVIAIGILAAALPAYODY 34

RESULT 10
US-08-485-264A-109
; Sequence 109, Application US/08485264A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,264A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-264A-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVIAIIGILAAIAPAYOEY 24
DB 11 IELMIVIAIIGILAAIAPAYODY 34

RESULT 11
US-08-474-349A-109
Sequence 109, Application US/08474349A
Patent No. 6333395
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
TITLE OF INVENTION: VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 517
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,349A
FILING DATE: 07-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-024
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-474-349A-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVIAIIGILAAIAPAYOEY 24
DB 11 IELMIVIAIIGILAAIAPAYODY 34

RESULT 12
US-08-470-896-109
Sequence 109, Application US/08470896
Patent No. 6479055
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,896
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-470-896-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVIAIIGILAAIAPAYOEY 24
DB 11 IELMIVIAIIGILAAIAPAYODY 34

```

RESULT 13
US-08-485-546A-109
; Sequence 109, Application US/08485546A
; Patent No. 6518013
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Peteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,546A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-546A-109

Query Match          92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. NO. 1,1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1  IELMIVIAIIGIILAAALPAYQYQ 24
      |||||:||||:||||:|
Db      11 IELMIVIAIIGIILAAALPAYQY 34

RESULT 14
US-08-487-266A-109
; Sequence 109, Application US/08487266A
; Patent No. 6824783
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Peteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; NUMBER OF SEQUENCES: 239

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,266A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-487-266A-109

Query Match          92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 IELMIVAIIGIILAAIAPAYOEY 24
Db      11 IELMIVAIIGIILAAVALPAYODY 34

RESULT 15
; Sequence 109, Application US/08484741
; Patent No. 6951717
GENERAL INFORMATION:
APPLICANT: Bolognesi, Danl P.
          Matthews, Thomas J.
          Wild, Carl T.
          Barney, Shawn O.
          Lambert, Dennis M.
          Petteway, Stephen R.
          Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
                     OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
                     TRANSMISSION
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,741
FILING DATE: 07-Jun-1995

```

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-022
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 109:
US-08-484-741-109

Query Match 92.2%; Score 106; DB 2; Length 167;
Best Local Similarity 87.5%; Pred. No. 1,1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IRLMIVIAIGILAVALPAYQY 24
| | | | | : | | | | | : | | | | | : |
Db 11 IRLMIVIAIGILAVALPAYQY 34

Search completed: January 23, 2006, 10:46:48
Job time : 23.9258 secs

The Page Blank (ustio)

NUMBER OF SEQ ID NOS: 763
SOFTWARE: PatentIn version 3.3
SEQ ID NO 238
LENGTH: 170
TYPE: PRT
ORGANISM: Neisseria meningitidis Z2491
US-11-052-554A-238

Query Match 92.2%; Score 106; DB 7; Length 170;
Best Local Similarity 87.5%; Pred. No. 1e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIIGILAIAPAYOEY 24
Db 11 IELMIVIAIIGILAIAPAYODY 34

RESULT 3
US-10-972-587-36
Sequence 36; Application US/10972587
Publication No. US20050246799A1
GENERAL INFORMATION:
APPLICANT: Song, Xiaojing
APPLICANT: Bartola, Pauline Anne
APPLICANT: Linderock, Nora Abiella
APPLICANT: Pan, Hao
TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 21829/213
CURRENT APPLICATION NUMBER: US/10/972,587
CURRENT FILING DATE: 2004-10-25
PRIOR APPLICATION NUMBER: 60/335,776
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: 09/810,997
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 10/174,209
PRIOR FILING DATE: 2002-06-17
NUMBER OF SEQ ID NOS: 86
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
LENGTH: 26
TYPE: PRT
ORGANISM: Xanthomonas campestris pv. glycines
US-10-972-587-36

Query Match 87.8%; Score 101; DB 6; Length 26;
Best Local Similarity 87.5%; Pred. No. 6.6e-09;
Matches 21; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IELMIVIAIIGILAIAPAYOEY 24
Db 3 IELMIVIAIIGILAIAPAYODY 26

RESULT 4
US-11-052-554A-82
Sequence 82; Application US/11052554A
Publication No. US20050288866A1
GENERAL INFORMATION:
APPLICANT: Sachdev, et al.
TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
FILE REFERENCE: 30853/40359A
CURRENT APPLICATION NUMBER: US/11/052,554A
CURRENT FILING DATE: 2005-02-07
PRIOR APPLICATION NUMBER: US 60/589,227
PRIOR FILING DATE: 2004-07-20
PRIOR APPLICATION NUMBER: IN 173/DEL/2004
PRIOR FILING DATE: 2004-02-06
NUMBER OF SEQ ID NOS: 763
SOFTWARE: PatentIn version 3.3
SEQ ID NO 82

LENGTH: 149
TYPE: PRT
ORGANISM: Haemophilus influenzae Rd
US-11-052-554A-82

Query Match 79.1%; Score 91; DB 7; Length 149;
Best Local Similarity 79.2%; Pred. No. 1.3e-06;
Matches 19; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IELMIVIAIIGILAIAPAYOEY 24
Db 16 IELMIVIAIIGILAIAPSYONY 39

RESULT 5
US-10-467-657-3494
Sequence 3494; Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 3494
LENGTH: 162
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3494

Query Match 74.8%; Score 86; DB 6; Length 162;
Best Local Similarity 56.0%; Pred. No. 7.2e-06;
Matches 14; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IELMIVIAIIGILAIAPAYOEY 25
Db 14 IELMIVIAIIGILAIAPSYQSYI 38

RESULT 6
US-10-467-657-6040
Sequence 6040; Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 6040
LENGTH: 129
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6040

Query Match 67.8%; Score 78; DB 6; Length 129;
Best Local Similarity 56.0%; Pred. No. 7.8e-05;

Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ELMIVIAIGIILAIAPAYQEV 25
:||||:||||:||||:||||:||||:
Db 11 ELMIVIAIGIILAIAPAYQEV 35

RESULT 7
US-10-467-657-688
; Sequence 688, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 688
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-688

Query Match 54.8%; Score 63; DB 6; Length 149;
Best Local Similarity 44.0%; Pred. No. 0.013;
Matches 11; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ELMIVIAIGIILAIAPAYQEV 25
:||||:||||:||||:||||:||||:
Db 10 VELISVLLISVALIIVPSRYNV 34

RESULT 8
US-10-467-657-3480
; Sequence 3480, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3480
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3480

Query Match 48.7%; Score 56; DB 6; Length 236;
Best Local Similarity 66.7%; Pred. No. 0.22;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELMIVIAIGIILAIAP 19
:||||:||||:||||:||||:||||:
Db 28 ELMIVIAIGIILAIAP 45

RESULT 9
US-10-878-556A-143
; Sequence 143, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21/62
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: sw_hum/call_human
; DATABASE ENTRY DATE: 1994-10-01
US-10-878-556A-143

Query Match 40.9%; Score 47; DB 6; Length 337;
Best Local Similarity 71.4%; Pred. No. 6.3;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 12 ILAIALPAYQEV 25
:||||:||||:||||:||||:||||:
Db 49 ILAALQMPAYQEV 62

RESULT 10
US-10-467-657-6830
; Sequence 6830, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6830
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6830

Query Match 38.3%; Score 44; DB 6; Length 223;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELMIVIAIGIILAIAP 18
:||||:||||:||||:||||:||||:
Db 149 ELPAVLLIGMVNIAL 165

RESULT 11
US-10-131-826A-224
; Sequence 224, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen

```

; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 224
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-224

Query Match      38.3%; Score 44; DB 6; Length 449;
Best Local Similarity 32.1%; Pred. No. 23;
Matches 9; Conservative 7; Mismatches 6; Indels 6; Gaps 1;

Qy      3 LMIYVIAIG-----ILAAIALPAYOEY 24
Db      401 LLLVLLGLGVTLFTVLVLPALQAYESY 428

RESULT 12
US-10-523-477-8
; Sequence 8, Application US/10523477
; Publication No. US2005026406A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MAXS AS MODIFIERS OF THE AXIN PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-051C-US
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US/10/523,477
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/401,534
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/411,153
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 519
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
US-10-523-477-8

Query Match      38.3%; Score 44; DB 6; Length 519;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy      6 VIALIGILAAIALP 19
Db      230 VIALIGILAAIALP 243

RESULT 13
US-10-821-234-1409
; Sequence 1409, Application US/10821234
; Publication No. US2005025511A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Strache-Crain, Blrgit
; APPLICANT: Andarmant, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PL SEQ_genes Version 1.0
; SEQ ID NO 1409
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1409

Query Match      38.3%; Score 44; DB 6; Length 645;
Best Local Similarity 44.4%; Pred. No. 35;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy      7 IATIGILAAIALPAYOEY 24
Db      306 VIIIGVFKEGSDPAYOEY 323

RESULT 14
US-11-198-847-143
; Sequence 143, Application US/11198847
; Publication No. US20050271589A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Olivera, Baldoemero M.
; TITLE OF INVENTION: B-Superfamily Conotoxins
; FILE REFERENCE: 2314-296
; CURRENT APPLICATION NUMBER: US/11/198,847
; CURRENT FILING DATE: 2005-08-08
; PRIOR APPLICATION NUMBER: US 10/838,226
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 10/058,053
; PRIOR FILING DATE: 2000-01-29
; PRIOR APPLICATION NUMBER: US 60/264323
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 143
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Conus miles
; FEATURE:
; NAME/KEY: misc_feature
```


/ LOCATION: (1)...(587)
/ OTHER INFORMATION: n may be any nucleotide
US-11-198-847-143

Query Match 37.4%; Score 43; DB 7; Length 109;
Best Local Similarity 28.6%; Pred. No. 6.6;
Matches 6; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 3 LMIYAIIGILAIAPAYOE 23
DB 8 MMDVMMGVGTWAGSLPFPDD 28

RESULT 15
US-10-517-939-330
/ Sequence 330, Application US/10517939
/ Publication No. US20060003433A1
/ GENERAL INFORMATION:
/ APPLICANT: Steer, Brian
/ APPLICANT: Callen, Walter
/ APPLICANT: Healey, Shaun
/ APPLICANT: Hazlewood, Geoff
/ APPLICANT: Wu, Di
/ APPLICANT: Blum, David
/ APPLICANT: Esteghlalian, Alireza
/ TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
/ TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
/ FILE REFERENCE: 564462007901
/ CURRENT APPLICATION NUMBER: US/10/517,939
/ CURRENT FILING DATE: 2004-12-13
/ PRIOR APPLICATION NUMBER: PCT/US03/19153
/ PRIOR FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: 60/389,299
/ PRIOR FILING DATE: 2002-06-14
/ NUMBER OF SEQ ID NOS: 360
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 330.
/ LENGTH: 755
/ TYPE: PRT
/ ORGANISM: Unknown
/ FEATURE:
/ OTHER INFORMATION: Obtained from an environmental sample.
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: (1)...(35)
US-10-517-939-330

Query Match 37.4%; Score 43; DB 6; Length 755;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 5 IYVIAIGILAIAPALPA 20
DB 17 LVIALPLLLAVALPS 32

Search completed: January 23, 2006, 12:09:41
Job time : 3.49345 secs

The Page Number
(1234567890)

CC stimulating factor (GM-CSF), enkephalin, dargirin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from p11n of
 CC Kingella dentrificans and is penetrating peptide 13 of the invention.
 XX
 XX
 SQ Sequence 25 AA;
 Query Match 100.0%; Score 115; DB 6; Length 25;
 Best Local Similarity 100.0%; Pred. No. 3.5e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TELMIVAIIGILAAIALPAYQERYV 25
 Db 1 TELMIVAIIGILAAIALPAYQERYV 25
 RESULT 2
 AEB08230
 ID AEB08230 standard; peptide; 25 AA.
 XX
 AC AEB08230;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DE Kingella dentrificans p11n penetrating peptide 13, SEQ ID NO: 13.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antipneumatic; cystostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Kingella dentrificans.
 XX
 PN US2005136103-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 16-SEP-2004; 2004US-00942300.
 XX
 PR 17-SEP-2003; 2003US-00664989.
 XX
 PR 17-SEP-2003; 2003US-00665184.
 XX
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 XX
 PA (COHEN/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 XX
 DR WPI, 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 46, SEQ ID NO 13; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Kingella dentrificans p11n
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 XX Sequence 25 AA;
 Query Match 100.0%; Score 115; DB 9; Length 25;
 Best Local Similarity 100.0%; Pred. No. 3.5e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TELMIVAIIGILAAIALPAYQERYV 25
 Db 1 TELMIVAIIGILAAIALPAYQERYV 25
 RESULT 3
 ADB16899
 ID ADB16899 standard; peptide; 25 AA.
 XX
 AC ADB16899;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Eikenella corrodens p11n penetrating peptide 14.
 XX
 KW penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 KW infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cystostatic; neurotropic.
 XX
 OS Eikenella corrodens.
 XX
 PN WO2003066859-A2.
 XX
 PD 14-AUG-2003.
 XX
 PD 07-FEB-2003; 2003WO-IB000968.
 XX
 PR 07-FEB-2002; 2002US-0355396P.
 XX
 PR (YISS) YISSUM RBS DEV CO HEBREW UNIV JERUSALEM.
 XX
 PA Ben-Sasson SA, Cohen E;
 XX
 DR WPI, 2003-697452/66.
 XX
 PT New penetrating peptide, useful for preparing a composition for treating

PT or preventing e.g. endocrine disorders.
 XX
 PS Claim 2, Page 14; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC antitoxinant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, hematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiatherosclerotic, osteopathic,
 CC cytoskeletal or neurotrophic activities. This peptide is from pillin of
 CC Bikenella corrodens and is penetrating peptide 14 of the invention.
 XX
 SQ Sequence 25 AA;
 Query Match 97.4%; Score 112; DB 6; Length 25;
 Best Local Similarity 96.0%; Pred. No. 1e-09; Indels 0; Gaps 0;
 Matches 24; Conservative 1; Mismatches 0;
 QY 1 IELMIVAIIGILAAIAPAYQRYV 25
 DB 1 IELMIVAIIGILAAIAPAYQRYV 25
 RESULT 4
 ID AEB08231
 AEB08231 standard; peptide; 25 AA.
 AC AEB08231;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DB Bikenella corrodens pillin penetrating peptide 14, SEQ ID NO: 14.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotrophic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotoxic;
 KW genitourinary disease; hematological disease; anti-anemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antitumefactory; hepatotropic; hepatitis B virus infection.
 XX
 OS Bikenella corrodens.
 XX
 PN US2005136103-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 16-SEP-2004; 2004US-00942300.

XX
 PR 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 DR WPI: 2005-444089/45.
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 14; 59pp; English.
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes;
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological;
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Bikenella corrodens pillin
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 SQ Sequence 25 AA;
 Query Match 97.4%; Score 112; DB 9; Length 25;
 Best Local Similarity 96.0%; Pred. No. 1e-09;
 Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IELMIVAIIGILAAIAPAYQRYV 25
 DB 1 IELMIVAIIGILAAIAPAYQRYV 25
 RESULT 5
 ID AAR38501
 AAR38501 standard; protein; 57 AA.
 XX
 AC AAR38501;
 XX
 DT 25-MAR-2003 (revised)
 DT 28-OCT-1993 (first entry)
 XX
 DE M. bovis pillin protein submolecular unit.
 XX
 KW Antibodies; whole pill binding; basis; vaccine; bacterial infection;
 KW tumulant footrot infection; sheep; type IV pilliated bacteria.
 XX
 OS Moraxella bovis.
 XX
 PN WO9311791-A1.
 XX
 PD 24-JUN-1993.

PR 05-OCT-2001; 2001US-0326998P.
 PR 07-NOV-2001; 2001US-0331045P.
 XX
 XX
 PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
 XX
 XX Renner WA, Bachmann M, Tisbot A, Maurer P, Lechner F, Sebbel P,
 PI Ploesek C,
 DR WPI; 2002-627351/67.
 XX
 XX Molecular antigen array used in the production of vaccines for infectious
 PT diseases.
 XX
 XX Disclosure; Page 370; 441pp; English.
 XX
 XX This invention relates to a novel ordered and repetitive antigen array
 CC used in the production of vaccines for infectious diseases. The invention
 CC also discloses a composition comprising a non-natural molecular scaffold
 CC comprising a core particle selected from a core particle of a non-natural
 CC origin and a core particle of natural origin and an organiser comprising
 CC at least one first attachment site, where the organiser is connected to
 CC the core particle by at least one covalent bond. Also disclosed is an
 CC antigen or antigenic determinant with at least one second attachment
 CC site, where the antigen or antigenic determinant is amyloid beta peptide
 CC (Abeta1-42) or its fragment and where the second attachment site is
 CC selected from an attachment site not naturally occurring with the antigen
 CC or antigenic determinant and an attachment site naturally occurring with
 CC the antigen or antigenic determinant, where the second attachment site is
 CC capable of association through at least one non-peptide bond to the first
 CC scaffold interact through the association to form an ordered and
 CC repetitive antigen array. The invention also comprises a coat protein
 CC capable of forming a capid which comprises mutant Obeta coat proteins
 CC having an amino acid sequence selected from five amino acid sequences
 CC fully defined in the specification. The compounds of the invention may
 CC have antimicrobial, antiallergic, immunomodulatory, cytoprotective,
 CC antiviral, antidiabetic, or hypoglycaemic activities and may be used in
 CC immunisation and as a vaccine. The present sequence represents a protein
 XX sequence used to create the compositions of the invention
 XX
 SQ Sequence 139 AA;
 XX
 Query Match 93.0%; Score 107; DB 5; Length 139;
 Best Local Similarity 91.7%; Pred. No. 3.9e-08;
 Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IELMIVAIIGIILAIAPAYQRY 24
 DB 12 IELMIVAIIGIILAIAPAYQRY 35
 XX
 XX RESULT 8
 XX ABG80534
 ID ABG80534 standard; protein; 139 AA.
 XX
 AC ABG80534;
 XX
 DT 29-NOV-2002 (first entry)
 XX
 DE Pseudomonas stutzeri pili protein.
 XX
 XX Molecular antigen array; vaccine; antigen; antimicrobial;
 CC molecular scaffold; amyloid beta; Abeta 1-42; influenza;
 CC graft versus host disease; IGE-mediated allergic reaction; anaphylaxis;
 CC adult respiratory distress syndrome; ARDS; Crohn's disease;
 CC allergic asthma; acute lymphoblastic leukaemia; non-Hodgkin's lymphoma;
 CC Grave's disease; systemic lupus erythematosus; osteoporosis;
 CC inflammatory immune disease; myasthenia gravis; multiple sclerosis;
 CC immunoproliferative disease; lymphadenopathy; Alzheimer's disease;
 CC angioimmunoproliferative lymphadenopathy; immunoblastic lymphadenopathy;
 CC rheumatoid arthritis; diabetes; infectious disease; factor Xa;
 KW enterokinase; cysteine-containing linker.
 XX

OS Pseudomonas stutzeri.
 XX
 XX WO200256907-A2.
 XX
 XX 25-JUL-2002.
 XX
 XX 21-JAN-2002; 2002WO-IB000168.
 XX
 XX 19-JAN-2001; 2001US-0262379P.
 PR 04-MAY-2001; 2001US-0286549P.
 PR 05-OCT-2001; 2001US-0326998P.
 PR 07-NOV-2001; 2001US-0331045P.
 XX
 XX (CYTO-) CYTOS BIOTECHNOLOGY AG.
 PA (NOVS) NOVARTIS PHARMA AG.
 PA (MAUR/) MAURER P.
 PA (LECH/) LECHNER F.
 PA (ORTM/) ORTMANN R.
 PA (LUEB/) LUEBEND R.
 PA (STAD/) STAUFENBIEL M.
 PA (FREY/) FREY P.
 XX
 XX Maurer P, Lechner F, Ortmann R, Luegend R, Staufenbiel M, Frey P,
 PI Renner WA, Bachmann M, Tisbot A, Sebbel P, Ploesek C,
 DR WPI; 2002-636514/68.
 XX
 XX Molecular antigen array used in the production of vaccines for infectious
 PT diseases.
 XX
 XX Disclosure; Page 347; 418pp; English.
 XX
 XX The invention relates to a composition comprising: (a) a non-natural
 CC molecular scaffold comprising: (1) a core particle selected from: (1) a
 CC core particle of a non-natural origin; and (2) a core particle of natural
 CC origin; and (1) an organiser comprising at least one first attachment
 CC site, where the organiser is connected to the core particle by at least
 CC one covalent bond; (b) an antigen or antigenic determinant with at least
 CC one second attachment site, where the antigen or antigenic determinant is
 CC amyloid beta peptide (Abeta 1-42) or its fragment, and where the second
 CC attachment site is selected from: (1) an attachment site not naturally
 CC occurring with the antigen or antigenic determinant; and (1) an
 CC attachment site naturally occurring with the antigen or antigenic
 CC determinant, where the second attachment site is capable of association
 CC through at least one non-peptide bond to the first attachment site; and
 CC where the antigen or antigenic determinant and the scaffold interact
 CC through the association to form an ordered and repetitive antigen array.
 CC Also included is a process for producing a non-naturally occurring
 CC ordered and repetitive antigen array. The composition is used in
 CC immunisation and as a vaccine for diseases such as influenza, graft
 CC versus host disease, IGE-mediated allergic reactions, anaphylaxis, adult
 CC respiratory distress syndrome (ARDS), Crohn's disease, allergic asthma,
 CC acute lymphoblastic leukaemia, non-Hodgkin's lymphoma, Grave's disease,
 CC systemic lupus erythematosus, inflammatory immune diseases, myasthenia
 CC gravis, immunoproliferative disease lymphadenopathy,
 CC angioimmunoproliferative lymphadenopathy, immunoblastic lymphadenopathy,
 CC rheumatoid arthritis, diabetes, multiple sclerosis, Alzheimer's disease,
 CC osteoporosis and infectious diseases. The antigens are modified to possess
 CC a cleavage site (enterokinase or factor Xa) and a Cysteine- containing N-
 CC or C-terminal linker peptide which serves as the attachment point to a
 CC present sequence is bacterial protein or peptide which is coupled to the
 CC modified antigen to form the molecular antigen array
 XX
 SQ Sequence 139 AA;
 XX
 Query Match 93.0%; Score 107; DB 5; Length 139;
 Best Local Similarity 91.7%; Pred. No. 3.9e-08;
 Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IELMIVAIIGIILAIAPAYQRY 24
 DB 12 IELMIVAIIGIILAIAPAYQRY 35

```

RESULT 9
ADD24109
ID ADD24109 standard; protein; 139 AA.
XX
AC ADD24109;
XX
DT 15-JAN-2004 (first entry)
XX
DE Pseudomonas stutzeri pili protein.
XX
KM vaccine composition; virus-like particle; core particle;
KM first attachment site; antigen; antigenic determinant; prion protein;
KM PrP peptide; vaccine; neuroprotective; antiinflammatory;
KM prion disease; Bovine Spongiform Encephalopathy; BSE;
KM Creutzfeldt-Jakob Disease; pili.
XX
OS Pseudomonas stutzeri.
XX
PN WO2003059386-A2.
XX
PD 24-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-EP000460.
XX
PR 18-JAN-2002; 2002US-00050902.
PR 21-JAN-2002; 2002WO-1B000166.
PR 08-JUL-2002; 2002US-0393725P.
PR 18-JUL-2002; 2002US-0396590P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
XX
PI Bachmann M, Maurer P, Pelliccioli E, Renner WA;
XX
DR WPI; 2003-598483/56.
XX
PT A vaccine composition for preventing or treating prion diseases (e.g.
PT Creutzfeldt-Jakob Disease) comprises a virus-like particle (e.g. RNA-
PT phage) and at least one prion protein or peptide bound to the virus-like
PT particle.
XX
PS Disclosure; SEQ ID NO 2; 246bp; English.
XX
SQ
CC This invention relates to a novel vaccine composition comprising a virus-
CC like or a core particle with at least one first attachment site and at
CC least one antigen or antigenic determinant that is a prion protein (PrP)
CC or its dimer, or a PrP peptide, the antigen or antigenic determinant
CC being bound to the virus-like or core particle. The vaccine of the
CC invention may have neuroprotective or antiinflammatory activity. The
CC composition is useful as a medicament or in manufacturing a medicament
CC for the treatment or prevention of prion diseases. The prion diseases may
CC include Bovine Spongiform Encephalopathy (BSE) or Creutzfeldt-Jakob
CC Disease. The present sequence is the amino acid sequence of the
CC Pseudomonas stutzeri pili protein which may be used during the creation
CC of the vaccine composition of the invention.
XX
SQ Sequence 139 AA;
Query Match 93.0%; Score 107; DB 7; Length 139;
Best Local Similarity 91.7%; Pred. No. 3.9e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 TELMTVIAIIGILAIALPAYQY 24
DB 12 TELMTVIAIIGILAIALPAYQY 35
RESULT 10
ADJ82035
ID ADJ82035 standard; protein; 139 AA.
XX
AC ADJ82035;

```

```

XX
DT 06-MAY-2004 (first entry)
XX
DE Protein for RANKL antigen array to treat bone disease.
XX
KM osteopathic; vaccine; core particle; antigenic determinant; RANKL;
KM bone disease; encephalopathy; immune system stimulation.
XX
OS Unidentified.
XX
PN WO2003039225-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-EP012449.
XX
PR 07-NOV-2001; 2001US-0331045P.
PR 18-JAN-2002; 2002US-00050902.
PR 21-JAN-2002; 2002WO-1B000166.
PR 19-JUL-2002; 2002US-0396635P.
XX
PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
XX
PI Bachmann M, Maurer P, Spohn G;
XX
DR WPI; 2003-441430/41.
XX
PT New compositions comprising a core particle and at least one antigen or
PT antigenic determinant, useful for as a vaccine for therapy or prophylaxis
PT of bone diseases, particularly mammalian encephalopathies.
XX
PS Disclosure; SEQ ID NO 2; 222bp; English.
XX
SQ
CC The invention relates to a composition comprising a core particle having
CC at least one first attachment site, and at least one antigen or antigenic
CC determinant having at least one second attachment site. The antigen or
CC antigenic determinant is a RANKL protein, RANKL fragment or RANKL
CC peptide. The second attachment site is (non-) naturally occurring with
CC the antigen or antigenic determinant, and is capable of association to
CC the first attachment site. The antigen or antigenic determinant and the
CC core particle interact through the association to form an ordered and
CC repetitive antigen array. The composition is useful as a medicament, or
CC for the manufacture of a medicament for treating bone diseases. The
CC composition is especially useful for as a vaccine for therapy or
CC prophylaxis of bone diseases, particularly mammalian encephalopathies,
CC and for stimulating mammalian immune system. This sequence represents a
CC protein of the invention.
XX
SQ Sequence 139 AA;
Query Match 93.0%; Score 107; DB 7; Length 139;
Best Local Similarity 91.7%; Pred. No. 3.9e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 TELMTVIAIIGILAIALPAYQY 24
DB 12 TELMTVIAIIGILAIALPAYQY 35
RESULT 11
ADK17123
ID ADK17123 standard; peptide; 139 AA.
XX
AC ADK17123;
XX
DT 06-MAY-2004 (first entry)
XX
DE Virus-like particle repetitive antigen array peptide #2.
XX
KM antiallergic; antiaesthetic; cytostatic; vaccine; virus-like particle;
KM interleukin; IL-5; IL-13; eotaxin; repetitive antigen array;
KM allergic eosinophilic disease; asthma; Hodgkin's lymphoma.
XX

```


OS Unidentified.
 XX MO2003040164-A2.
 PN 15-MAY-2003.
 PD 07-NOV-2002; 2002MO-BP012455.
 PF 07-NOV-2001; 2001US-0331045P.
 PR 18-JAN-2002; 2002US-00050902.
 PR 21-JAN-2002; 2002MO-1B000166.
 PR 19-JUL-2002; 2002US-0396636P.
 XX (CYTO-) CYTOS BIOTECHNOLOGY AG.
 PA Bachmann M, Jennings G, Sonderegger I;
 PI WPI; 2003-441518/41.
 DR Composition comprising an ordered and repetitive antigen or antigenic
 XX determinant array, useful as a medicament, or for manufacturing a
 PT medicament for treating allergic eosinophilic diseases, e.g. asthma, or
 PT Hodgkin's lymphoma.
 PS Disclosure; SEQ ID NO 2; 245bp; English.
 XX The invention relates to a composition comprising a virus-like particle
 CC and at least one antigen, which is a protein or peptide of interleukin
 CC (IL)-5, IL-13 or eotaxin and is bound to the virus-like particle, or a
 CC core particle with at least one first attachment site and at least one
 CC antigen with at least one second attachment site, where the antigen is a
 CC protein or peptide of IL-5, IL-13 or eotaxin. The second attachment site
 CC is an attachment site naturally or not naturally occurring with the
 CC antigen or antigenic determinant. The second attachment site is capable
 CC of association to the first attachment site, and where the antigen or
 CC antigenic determinant and the core particle interact through the
 CC association to form an ordered and repetitive antigen array. The
 CC compositions are useful as medicaments, or for manufacturing a medicament
 CC or a vaccine for treating allergic eosinophilic diseases, e.g. asthma, or
 CC Hodgkin's lymphoma and related diseases. This sequence is used to
 CC generate the compound of the invention.
 SQ Sequence 139 AA;
 Query Match 93.0%; Score 107; DB 7; Length 139;
 Best Local Similarity 91.7%; Pred. No. 3.9e-08;
 Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IELMTVIAITIGILAAIALPAYQEV 24
 DB 12 IELMTVIAITIGILAAIALPAYQDY 35
 RESULT 12
 ADA34990
 ID ADA34990 standard; protein; 169 AA.
 AC ADA34990;
 DT 20-NOV-2003 (first entry)
 XX Acinetobacter baumannii protein #2151.
 DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.
 OS Acinetobacter baumannii.
 XX US6562958-B1.
 PN 13-MAY-2003.
 PD 04-JUN-1999; 99US-00328352.
 PF

XX 09-JUN-1998; 98US-0086701P.
 PR (GENO-) GENOME THERAPEUTICS CORP.
 XX Breton G, Bush D;
 PI WPI; 2003-576092/54.
 DR N-PSDB; ADA30864.
 XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 PS Example; SEQ ID NO 6277; 328bp; English.
 XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 SQ Sequence 169 AA;
 Query Match 93.0%; Score 107; DB 6; Length 169;
 Best Local Similarity 88.0%; Pred. No. 4.8e-08;
 Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 IELMTVIAITIGILAAIALPAYQEV 25
 DB 26 IELMTVIAITIGILAAIALPAYQNV 50
 RESULT 13
 AAP40503
 ID AAP40503 standard; protein; 50 AA.
 AC AAP40503;
 DT 25-MAR-2003 (revised)
 DT 31-JAN-1992 (first entry)
 XX Sequence of the CNI fragment (residues 1-50) of pill from strain F62 of
 DE Neisseria gonorrhoeae.
 XX Neisseria gonorrhoeae vaccine; purified pill.
 KW Neisseria gonorrhoeae.
 OS Neisseria gonorrhoeae.
 XX Key Location/Qualifiers
 FH Region 1..7
 FT /label= CNI11
 FT Modified-site 1
 FT /label= N-Me-Phe
 FT Region 8..50
 FT /label= CNI
 XX US4443431-A.
 PN 17-APR-1984.
 PD 27-MAY-1981; 81US-00267538.
 PF 27-MAY-1981; 81US-00267538.
 XX (USSH) US SEC DEPT HEALTH.
 XX Buchanan TM, Pearce W, Chen KCS;
 PI

XX WPI; 1984-113477/18.
XX
XX Neisseria gonorrhoeae vaccine - contg. pili fragments instead of whole
PT pili.
XX
XX Example; Table 2, column 5 and 6; 6pp; English.
XX
XX To create the fragments of pili, pili purified according to methods
CC described in Pearce and Buchanan (1978), or Hermodson (1978) were
CC reduced, carboxymethylated and then cleaved using cyanogen bromide. The
CC largest fragment, comprising approximately 60% of the subunit of strain
CC F62 N.9. pili, was termed CNI. The second fragment peak, approximately
CC 32% of the subunit, was termed CNII, and a final small fragment of only 7
CC AAs was termed CNIII. A novel Neisseria gonorrhoeae vaccine consists of a
CC CNI fragment. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 50 AA;

Query Match 92.2%; Score 106; DB 1; Length 50;
Best Local Similarity 87.5%; Pred. No. 1.8e-08; Indels 0; Gaps 0;
Matches 21; Conservative 3; Mismatches 0;

QY 1 IELMIVIAIGILAAALPAYQY 24
|||||:|||||:|||||:
DB 4 IELMIVIAIGILAAALPAYQDY 27

RESULT 14

AAR38500
ID AAR38500 standard; protein; 53 AA.
XX

AC AAR38500;
XX

DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-OCT-1993 (first entry)
XX

DE N. meningitis pili protein submolecular unit.
XX

KW Antibodies; whole pili binding; basis; vaccine; bacterial infection;
KW ruminant footrot infection; sheep; type IV pilated bacteria.
XX

OS Neisseria meningitidis.
XX

PN W09311791-A1.
XX

PD 24-JUN-1993.
XX

PP 17-DEC-1992; 92WO-US011085.
XX

PR 18-DEC-1991; 91US-00809762.
XX

PA (UYOR-) UNIV OREGON HEALTH SCI.
XX

PI Smith AW;
XX

DR WPI; 1993-213824/26.
XX

PT Antigenic preparation - stimulates production of antibodies binding to
PT pili protein of type IV pilated bacteria, useful in vaccine compn.
XX

PS Claim 9; Page 25; 44pp; English.
XX

CC The sequence is that of a submolecular unit of Neisseria meningitis pili
CC protein which corresponds to at least one epitope common to structural
CC pili proteins of type IV pilated bacteria. It is capable of eliciting
CC antibodies (Abs) which bind to whole pili of type IV bacteria. The
CC ability of the submolecular unit to produce Abs which bind to whole pili
CC provides the basis for vaccines against type IV bacterial infections,
CC e.g. footrot infection in ruminants. (Updated on 25-MAR-2003 to correct
CC PN field.) (Updated on 27-AUG-2003 to correct OS field.)
XX

SQ Sequence 53 AA;

Query Match 92.2%; Score 106; DB 2; Length 53;
Best Local Similarity 87.5%; Pred. No. 1.9e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVIAIGILAAALPAYQY 24
|||||:|||||:|||||:
DB 4 IELMIVIAIGILAAALPAYQDY 27

RESULT 15

AAR38499
ID AAR38499 standard; protein; 53 AA.
XX

AC AAR38499;
XX

DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-OCT-1993 (first entry)
XX

DE N. gonorrhea pili protein submolecular unit.
XX

KW Antibodies; whole pili binding; basis; vaccine; bacterial infection;
KW ruminant footrot infection; sheep; type IV pilated bacteria.
XX

OS Neisseria gonorrhoeae.
XX

PN W09311791-A1.
XX

PD 24-JUN-1993.
XX

PP 17-DEC-1992; 92WO-US011085.
XX

PR 18-DEC-1991; 91US-00809762.
XX

PA (UYOR-) UNIV OREGON HEALTH SCI.
XX

PI Smith AW;
XX

DR WPI; 1993-213824/26.
XX

PT Antigenic preparation - stimulates production of antibodies binding to
PT pili protein of type IV pilated bacteria, useful in vaccine compn.
XX

PS Claim 6; Page 25; 44pp; English.
XX

CC The sequence is that of a submolecular unit of Neisseria gonorrhea pili
CC protein which corresponds to at least one epitope common to structural
CC pili proteins of type IV pilated bacteria. It is capable of eliciting
CC antibodies (Abs) which bind to whole pili of type IV bacteria. The
CC ability of the submolecular unit to produce Abs which bind to whole pili
CC provides the basis for vaccines against type IV bacterial infections,
CC e.g. footrot infection in ruminants. (Updated on 25-MAR-2003 to correct
CC PN field.) (Updated on 27-AUG-2003 to correct OS field.)
XX

SQ Sequence 53 AA;

Query Match 92.2%; Score 106; DB 2; Length 53;
Best Local Similarity 87.5%; Pred. No. 1.9e-08;
Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVIAIGILAAALPAYQY 24
|||||:|||||:|||||:
DB 4 IELMIVIAIGILAAALPAYQDY 27

Search completed: January 23, 2006, 10:28:30
Job time : 84.9694 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OW protein - protein search, using sw model

Run on: January 23, 2006, 09:31:19 / Search time 79.6943 Seconds
(without alignments)
221.323 Million cell updates/sec

Title: US-10-501-838a-13
Perfect score: 115
Sequence: 1 IELMTVAITIGILAIAPAYQEVY 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_05.80:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	156	2	P77876_9NEIS
2	112	97.4	153	1	ECPC_EIKCO
3	111	96.5	154	1	FMW_MORBO
4	111	96.5	156	2	O59507_MORBO
5	111	96.5	157	1	FMQ_MORBO
6	111	96.5	157	2	O59503_MORBO
7	111	96.5	158	2	O59501_MORBO
8	111	96.5	158	2	O59508_MORBO
9	111	96.5	159	1	FMI_MORBO
10	111	96.5	159	2	O60123_MORBO
11	111	96.5	159	2	O59505_MORBO
12	111	96.5	160	2	O59504_MORBO
13	109	94.8	158	2	O8EJPS_SHEON
14	108	93.9	144	2	O9R4K1_EIKCO
15	108	93.9	144	2	P78000_9NEIS
16	108	93.9	153	2	O85826_EIKCO
17	107	93.0	139	2	O9ZEL3_PSEST
18	107	93.0	155	2	O8KQJ2_PSEAR
19	107	93.0	178	2	O8KQJ6_PSEAR
20	106	92.2	144	2	O5P985_AZOSR
21	106	92.2	148	2	O31392_NEIME
22	106	92.2	148	2	O9L9W1_NEILA
23	106	92.2	152	1	ECPD_EIKCO
24	106	92.2	159	1	ECPA_EIKCO
25	106	92.2	159	2	O51321_NEIGO
26	106	92.2	160	2	O51234_NEIME
27	106	92.2	160	2	O51235_NEIME
28	106	92.2	160	2	O57135_NEIME
29	106	92.2	163	2	O51236_NEIME
30	106	92.2	163	2	O51237_NEIME
31	106	92.2	163	2	O57441_NEIME

32	106	92.2	164	2	O52651_NEIGO
33	106	92.2	164	2	O8RLT8_NEIME
34	106	92.2	165	1	FMW_NEIGO
35	106	92.2	165	2	O52652_NEIGO
36	106	92.2	165	2	O54518_NEIGO
37	106	92.2	165	2	P77910_NEIGO
38	106	92.2	165	2	P77911_NEIGO
39	106	92.2	166	2	O08366_NEIME
40	106	92.2	166	2	O59621_NEIME
41	106	92.2	167	1	FMW_NEIGO
42	106	92.2	167	2	O52649_NEIGO
43	106	92.2	167	2	O8RLT8_NEIME
44	106	92.2	167	2	O59611_NEIGO
45	106	92.2	168	2	O59600_NEIGO
					O52651_neisseria g
					O8RLT8_neisseria m
					P02874_neisseria g
					O52652_neisseria g
					O54518_neisseria g
					P77910_neisseria g
					P77911_neisseria g
					O08366_neisseria m
					O59621_neisseria m
					P11764_neisseria g
					O52649_neisseria g
					O8RLT8_neisseria m
					O59611_neisseria g
					O59600_neisseria g

ALIGNMENTS

RESULT 1
P77876_9NEIS PRELIMINARY; PRT; 156 AA.
ID P77876.1
AC P77876.1
DT 01-FEB-1997 (TrEMBLrel_02, Created)
DT 01-FEB-1997 (TrEMBLrel_02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel_25, Last annotation update)
DE Kdgd.
GN Name=kdpd;
OS Kingella denitrificans.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Kingella.
ON NCBI_TaxID=502;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33394;
RX MEDLINE=97101010; PubMed=8945537;
RA Weir S., Lee L.W., Marrs C.P.;
RT "Identification of four complete type 4 pilin genes in a single
RT Kingella denitrificans genome."
RL Infect. Immun. 64:4993-4999 (1996).
DR EMBL: U68761; AAB48594.1; -, Genomic DNA.
DR HSP: P02974; 2PIL.
DR GO: GO:0009289; C: Cytodermis; IEA.
DR GO: GO:0015627; C: Cytotype II protein secretion system complex; IEA.
DR GO: GO:0008565; F: Protein transporter activity; IEA.
DR GO: GO:0007155; P: Cell adhesion; IEA.
DR GO: GO:0015628; P: Type II protein secretion system; IEA.
DR InterPro: IPR002416; Bac_GSPH.
DR InterPro: IPR001082; Pilin.
DR InterPro: IPR001120; Pili_N_methyl_S.
DR Pfam: PF00114; Pilin; 1.
DR PRINTS: PR00885; BACTERIALGSPH.
DR PRODOM: PD000666; Pilin; 1.
DR PROSITE: PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria.
SQ SEQUENCE 156 AA; 16766 MW; BE050AFED056A39 CRC64;
Query Match 100.0%; Score 115; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IELMTVAITIGILAIAPAYQEVY 25
Db 15 IELMTVAITIGILAIAPAYQEVY 39
RESULT 2
ECPC_EIKCO STANDARD; PRT; 153 AA.
ID ECPC_EIKCO
AC O07564;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)

```

DE Fimbrial protein ecpC precursor (Pilin).
GN Name=ecpC;
OS Eikenella corrodens.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Eikenella.
CX NCBI_TaxID=539;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=31745;
RX MEDLINE=93239296; PubMed=8478080;
RA Tonjum T., Weir S., Boye K., Proglake-Fox A., Marrs C.F.;
RT "Sequence divergence in two tandemly located pilin genes of Eikenella
corrodens.";
RL Infect. Immun. 61:1909-1916(1993).
CC -1- SIMILARITY: Belongs to the N-Me-Phe pil family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
CC EMBL, L12049; AAA03016.1; -; Unassigned_DNA.
DR HSSP; P02974; 2PIL.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR012902; Methylation_N.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF07963; N_methyl_1.
DR Pfam; PF00114; Pilin; 1.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
DR FIMBRIA; Methylation.
FT PROPEP 1 8 Potential.
FT CHAIN 9 153 Fimbrial protein ecpC.
FT MOD RES 9 9 N-methylphenylalanine (By similarity).
FT DISTUFD 130 143 By similarity.
SQ SEQUENCE 153 AA; 15695 MW; 03A98B1C8AFAA053 CRC64;

Query Match 97.4%; Score 112; DB 1; Length 153;
Best Local Similarity 96.0%; Pred. No. 5e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTILAIAPAYQEV 25
DB 12 IELMTVIAIIGTILAIAPAYQDYV 36

RESULT 3
FMM_MORNO STANDARD; PRT; 154 AA.
ID FMM_MORNO
AC P09829;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Fimbrial protein precursor (Pilin).
GN Name=tipA;
OS Moraxella nonliquefaciens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
CX NCBI_TaxID=478;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCTC 7784;
RX MEDLINE=92121902; PubMed=1770363;
RA Tonjum T., Marrs C.F., Rozsa F.W., Boye K.;
RT "The type 4 pilin of Moraxella nonliquefaciens exhibits unique
similarity with the pilins of Neisseria gonorrhoeae and
Dichelobacter (Bacteroides) nodosus.";
RL J. Gen. Microbiol. 137:2483-2490(1991).
RN [2]
RP PROTEIN SEQUENCE OF 7-55.

```

```

RX MEDLINE=77116125; PubMed=838045; DOI=10.1016/0014-5793(77)80008-2;
RA Frohnm L.O., Sletten K.;
RT "Purification and N-terminal sequence of a fimbrial protein from
Moraxella nonliquefaciens.";
RL FEBS Lett. 73:29-32(1977).
CC -1- SUBUNIT: The pil is a polar flexible filaments of about 5.4
nanometers diameter and 2.5 micrometers average length; they
consist of only a single polypeptide chain arranged in a helical
configuration of five subunits per turn in the assembled pilus.
CC -1- SIMILARITY: Belongs to the N-Me-Phe pil family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
CC EMBL, M59707; AAA25310.1; -; Genomic_DNA.
DR PIR; A44809; A44809.
DR HSSP; P02974; 2PIL.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR012902; Methylation_N.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF07963; N_methyl_1.
DR Pfam; PF00114; Pilin; 1.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
DR KW Direct protein sequencing; Fimbria; Methylation.
FT PROPEP 1 6
FT CHAIN 7 154 Fimbrial protein.
FT MOD RES 7 7 N-methylphenylalanine.
SQ SEQUENCE 154 AA; 16467 MW; 80BE4B0815563B0F CRC64;

Query Match 96.5%; Score 111; DB 1; Length 154;
Best Local Similarity 92.0%; Pred. No. 6.9e-07;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTILAIAPAYQEV 25
DB 10 IELMTVIAIIGTILAIAPAYQDYI 34

RESULT 4
Q59507_MORBO PRELIMINARY; PRT; 156 AA.
ID Q59507;
AC Q59507;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pilin.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
CX NCBI_TaxID=476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90256243; PubMed=1971258;
RA Elleman T.C., Hoyne P.A., Lepper A.W.;
RT "Characterization of the pilin gene of Moraxella bovis Dalton 2d and
expression of pil from M. bovis in Pseudomonas aeruginosa.";
RL Infect. Immun. 58:1678-1684(1990).
CC EMBL, M92155; AAA53087.1; -; Genomic_DNA.
DR PIR; A41490; A41490.
DR HSSP; P02973; 1OQM.
DR SMR; Q59507; 7-135.
DR GO; GO:0009289; C:fimbrium; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR002416; Bac_GSPH.

```

```

DR InterPro; IPR001082; PiliN.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF00114; PiliN; 1.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; PiliN; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
DR Fimbrin.
SQ SEQUENCE 156 AA; 15941 MW; EB281DA63B9B3C9 CRC64;

Query Match 96.5%; Score 111; DB 2; Length 156;
Best Local Similarity 92.0%; Pred. No. 6; 9e-07;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IELMIVIAITIGILAAIALPAYQEVY 25
DB 10 IELMIVIAITIGILAAIALPAYQDYI 34

RESULT 5
FMO MORBO STANDARD; PRT; 157 AA.
AC P07640;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Fimbrin protein Q precursor (Beta piliN) (Q piliN).
GN Name: fmo;
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxId=476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRP63;
RA MEDLINE=85234350; PubMed=2861194;
RA Martz C.F., Schoolnik G., Komey J.M., Hardy J., Rochard J.,
RA Falkow S.;
RT "Cloning and sequencing of a Moraxella bovis piliN gene.";
RL J. Bacteriol. 163:132-139(1985).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BRP63;
RA MEDLINE=90094235; PubMed=2403542;
RA Pulke K.A., Martz C.F., Stevens S.P., Green M.R.;
RT "Sequence analysis of the inversion region containing the piliN genes
RT of Moraxella bovis";
RL J. Bacteriol. 172:310-316(1990).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=91286182; PubMed=2061282;
RA Bozza P.W., Martz C.F.;
RT "Interesting sequence differences between the piliN gene inversion
RT regions of Moraxella lacunata ATCC 17956 and Moraxella bovis BRP63.";
RL J. Bacteriol. 173:400-406(1991).
RN [4]
RP PROTEIN SEQUENCE OF 7-157.
RA MEDLINE=89010522; PubMed=2902184; DOI=10.1084/jem.168.3.983;
RA Rueli W.W., Martz C.F., Fernandez R., Falkow S., Schoolnik G.K.;
RT "Purification, characterization, and pathogenicity of Moraxella bovis
RT piliN";
RL J. Exp. Med. 168:983-1002(1988).
CC -1- SUBUNIT: The pili are polar flexible filaments of about 5-4
CC nanometers diameter and 2.5 micrometers average length; they
CC consist of only a single polypeptide chain arranged in a helical
CC configuration of five subunits per turn in the assembled pilus.
CC -1- MISCELLANEOUS: Moraxella bovis can express either a Q or a I
CC piliN, the inversion of 2 kb of DNA determines which piliN is
CC expressed.
CC -1- SIMILARITY: Belongs to the N-Me-Phe pili family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; M11435; AAA25304.1; -; Genomic DNA.
CC DR EMBL; M32345; AAA88223.1; -; Genomic DNA.
CC DR EMBL; M59712; AAA25308.1; -; Genomic DNA.
CC DR PIR; A24434; A24434.
CC DR HSSP; P02973; 100W.
CC DR InterPro; IPR000983; Bac_GSPG.
CC DR InterPro; IPR002416; Bac_GSPH.
CC DR InterPro; IPR012902; Methylation_N.
CC DR InterPro; IPR001082; PiliN.
CC DR InterPro; IPR001120; Prok_N_methyl_S.
CC DR Pfam; PF07963; N_methyl_1.
CC DR Pfam; PF00114; PiliN; 1.
CC DR PRINTS; PR00813; BCTERIALGSPG.
CC DR PRINTS; PR00885; BCTERIALGSPH.
CC DR PRODOM; PD000666; PiliN; 1.
CC DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
CC DR Fimbrin.
CC KW Direct protein sequencing; Fimbrin; Methylation.
CC FT PROPEP 1 157
CC FT CHAIN 7 157 Fimbrin protein Q.
CC FT MOD RES 7 7 N-methylphenylalanine.
CC FT DIST:PID 136 155 By similarity.
CC SQ SEQUENCE 157 AA; 16006 MW; A923CDBA26C693C9 CRC64;

Query Match 96.5%; Score 111; DB 1; Length 157;
Best Local Similarity 92.0%; Pred. No. 7e-07;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IELMIVIAITIGILAAIALPAYQEVY 25
DB 10 IELMIVIAITIGILAAIALPAYQDYI 34

RESULT 6
ID Q59503 MORBO PRELIMINARY; PRT; 157 AA.
AC Q59503;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE PrepiliN.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxId=476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FL462 serogroup G;
RA MEDLINE=94327452; PubMed=8051000;
RA Atwell J.L., Tennent J.M., Lepper A.W., Elleman T.C.;
RT "Characterization of piliN genes from seven serologically defined
RT prototype strains of Moraxella bovis";
RL J. Bacteriol. 176:4875-4882(1994).
RN [2]
RA EMBL; L32966; AAA5359.1; -; Genomic DNA.
CC DR HSSP; P02973; 100W.
CC DR GO; GO:0009289; C:fimbrium; IEA.
CC DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
CC DR GO; GO:0008665; F:protein transporter activity; IEA.
CC DR GO; GO:0007155; P:cell adhesion; IEA.
CC DR GO; GO:0015628; P:cell protein secretion system; IEA.
CC DR InterPro; IPR000983; Bac_GSPG.
CC DR InterPro; IPR002416; Bac_GSPH.
CC DR InterPro; IPR001082; PiliN.
CC DR InterPro; IPR001120; Prok_N_methyl_S.
CC DR Pfam; PF00114; PiliN; 1.
CC DR PRINTS; PR00813; BCTERIALGSPG.
CC DR PRINTS; PR00885; BCTERIALGSPH.
CC DR PRODOM; PD000666; PiliN; 1.
CC DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
CC DR Fimbrin.
CC SQ SEQUENCE 157 AA; 16327 MW; EB4FCA0F382430D2 CRC64;

```

Query Match 96.5%; Score 111; DB 2; Length 157;
 Best Local Similarity 92.0%; Pred. No. 7e-07;
 Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTILAIAPAYQEVY 25
 |||||
 DB 10 IELMTVIAIIGTILAIAPAYQDYI 34

RESULT 7
 CS9501 MORBO PRELIMINARY; PRT; 158 AA.
 ID 059501 MORBO
 AC 059501
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Prepilin.
 OS Moraxella bovis.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 CC Moraxellaceae; Moraxella.
 OK NCBI_TaxID=476;
 RN Fimbria.
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=218 serogroup F;
 RX MEDLINE=94327452; PubMed=8051000;
 RA Atwell J.L., Tement J.M., Lepper A.W., Elleman T.C.;
 RT "Characterization of pilin genes from seven serologically defined
 RT prototype strains of Moraxella bovis.";
 RT J. Bacteriol. 176:4875-4882(1994).
 DR EMBL, U32965; AAA53558.1; -; Genomic_DNA.
 DR PIR, A55851; A55851.
 DR HSSP, P02973; 100W.
 DR GO, GO:0009289; C:fimbrium; IEA.
 DR GO, GO:0015627; C:type II protein secretion system complex; IEA.
 DR GO, GO:0008565; F:protein transporter activity; IEA.
 DR GO, GO:0007155; P:cell adhesion; IEA.
 DR GO, GO:0015628; P:type II protein secretion system; IEA.
 DR InterPro, IPR00983; Bac_GSPG.
 DR InterPro, IPR002416; Bac_GSPH.
 DR InterPro, IPR001082; Pilin.
 DR InterPro, IPR001120; Prok_N_methyl_S.
 DR Pfam, PF00114; Pilin; 1.
 DR PRINTS, PR00813; BCTERIALGSPG.
 DR PRINTS, PR00885; BCTERIALGSPH.
 DR PRODOM, PD000666; Pilin; 1.
 DR PROSITE, PS00409; PROKAR_NTER_METHYL, 1.
 KW Fimbria.
 SQ SEQUENCE 158 AA; 16216 MW; 620B513CA4DC39A5 CRC64;

Query Match 96.5%; Score 111; DB 2; Length 158;
 Best Local Similarity 92.0%; Pred. No. 7e-07;
 Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTILAIAPAYQEVY 25
 |||||
 DB 10 IELMTVIAIIGTILAIAPAYQDYI 34

RESULT 8
 CS9508 MORBO PRELIMINARY; PRT; 158 AA.
 ID 059508 MORBO
 AC 059508
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Prepilin.
 OS Moraxella bovis.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 CC Moraxellaceae; Moraxella.
 OK NCBI_TaxID=476;
 RN NUCLEOTIDE SEQUENCE.

RC STRAIN=3W07 serogroup B;
 RX MEDLINE=94327452; PubMed=8051000;
 RA Atwell J.L., Tement J.M., Lepper A.W., Elleman T.C.;
 RT "Characterization of pilin genes from seven serologically defined
 RT prototype strains of Moraxella bovis.";
 RT J. Bacteriol. 176:4875-4882(1994).
 DR EMBL, U32969; AAA53565.1; -; Genomic_DNA.
 DR HSSP, P02973; 100W.
 DR GO, GO:0009289; C:fimbrium; IEA.
 DR GO, GO:0015627; C:type II protein secretion system complex; IEA.
 DR GO, GO:0008565; F:protein transporter activity; IEA.
 DR GO, GO:0007155; P:cell adhesion; IEA.
 DR GO, GO:0015628; P:type II protein secretion system; IEA.
 DR InterPro, IPR00983; Bac_GSPG.
 DR InterPro, IPR002416; Bac_GSPH.
 DR InterPro, IPR001082; Pilin.
 DR InterPro, IPR001120; Prok_N_methyl_S.
 DR Pfam, PF00114; Pilin; 1.
 DR PRINTS, PR00813; BCTERIALGSPG.
 DR PRINTS, PR00885; BCTERIALGSPH.
 DR PRODOM, PD000666; Pilin; 1.
 DR PROSITE, PS00409; PROKAR_NTER_METHYL, 1.
 KW Fimbria.
 SQ SEQUENCE 158 AA; 16590 MW; 5958DC65678051F6 CRC64;

Query Match 96.5%; Score 111; DB 2; Length 158;
 Best Local Similarity 92.0%; Pred. No. 7e-07;
 Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTILAIAPAYQEVY 25
 |||||
 DB 10 IELMTVIAIIGTILAIAPAYQDYI 34

RESULT 9
 FMI MORBO STANDARD; PRT; 159 AA.
 ID FMI MORBO
 AC P20657;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Fimbrial protein I precursor (Alpha pilin) (I pilin).
 OS Moraxella bovis.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 CC Moraxellaceae; Moraxella.
 OK NCBI_TaxID=476;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=EP63;
 RC MEDLINE=90094235; PubMed=2403542;
 RX Fulk K.A., Marrs C.F., Stevens S.P., Green M.R.;
 RT "Sequence analysis of the inversion region containing the pilin genes
 RT of Moraxella bovis.";
 RT J. Bacteriol. 172:310-316(1990).
 [2]
 PROTEIN SEQUENCE OF 7-159.
 -1- SUBUNIT: The pilI are polar flexible filaments of about 5.4
 nanometers diameter and 2.5 micrometers average length; they
 consist of only a single polypeptide chain arranged in a helical
 configuration of five subunits per turn in the assembled pilus.
 -1- MISCELLANEOUS: Moraxella bovis can express either a Q or a I
 pilin, the inversion of 2 kb of DNA determines which pilin is
 expressed.
 -1- SIMILARITY: Belongs to the N-Me-Phe pilI family.

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use as long as its content is in no way modified and this statement is not removed.

CC -----

CC EMBL: M32345; -; NOT_ANNOTATED_CDS; Genomic_DNA.

DR FIR: J10071; J10071.

DR HSP: P02973; 100W.

DR InterPro: IPR002416; Bac_GSPH.

DR InterPro: IPR012902; Methylatlon_N.

DR InterPro: IPR001082; Piliin.

DR InterPro: IPR001120; Prok_N_methyl_S.

DR Pfam: PF07963; N_methyl_1.

DR Pfam: PF00114; Piliin; 1.

DR PRINTS: PR00885; BCTERIALGSPH.

DR ProDom: PD000666; Piliin; 1.

DR PROSITE: PS00409; PROKAR_NTER_METHYL; 1.

KM Direct protein sequencing; Fimbria; Methylatlon.

FT PROSP 1 6 Fimbrial protein I.

FT CHAIN 7 159 N-methylphenylalanine.

FT MOD RES 7 159 K -> KSK (in Ref. 2).

FT CONFLICT 159 159 K -> KSK (in Ref. 2).

SO SEQUENCE 159 AA; 16723 MW; 9130E2289C7F679E CRC64;

Query Match 96.5%; Score 111; DB 1; Length 159;

Best Local Similarity 92.0%; Pred. No. 7e-07;

Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAITIGILAAIALPAYQDYI 25

DB 10 IELMTVIAITIGILAAIALPAYQDYI 34

RESULT 10

Q60163 MORBO PRELIMINARY; PRT; 159 AA.

ID Q60163 MORBO PRELIMINARY; PRT; 159 AA.

AC Q60163;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)

DE Prepiliin.

OS Moraxella bovis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Moraxellaceae; Moraxella.

OX NCBI_TaxId=476;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=RS93L serotype D, and H358CS serogroup D;

RA MEDLINE=94327452; PubMed=8051000;

RT "Characterization of piliin genes from seven serologically defined RT prototype strains of Moraxella bovis."

RL J. Bacteriol. 176:4875-4882(1994).

EMBL: L32971; AAAS3563.1; -; Genomic_DNA.

EMBL: L32970; AAAS3562.1; -; Genomic_DNA.

DR HSP: P02973; 100W.

DR GO: GO:0009289; C:fimbrium; IEA.

DR GO: GO:0015627; C:type II protein secretion system complex; IEA.

DR GO: GO:0008565; F:protein transporter activity; IEA.

DR GO: GO:0007155; P:cell adhesion; IEA.

DR GO: GO:0015628; P:type II protein secretion system; IEA.

DR InterPro: IPR000983; Bac_GSPH.

DR InterPro: IPR002416; Piliin.

DR InterPro: IPR001082; Piliin.

DR Pfam: PF00114; Piliin; 1.

DR PRINTS: PR00813; BCTERIALGSPH.

DR ProDom: PD000666; Piliin; 1.

DR PROSITE: PS00409; PROKAR_NTER_METHYL; 1.

KM Fimbria.

SO SEQUENCE 159 AA; 16459 MW; F211DD7ECDS578FC CRC64;

Query Match 96.5%; Score 111; DB 2; Length 159;

Best Local Similarity 92.0%; Pred. No. 7e-07;

Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAITIGILAAIALPAYQDYI 25

DB 10 IELMTVIAITIGILAAIALPAYQDYI 34

RESULT 11

Q59505 MORBO PRELIMINARY; PRT; 159 AA.

ID Q59505 MORBO PRELIMINARY; PRT; 159 AA.

AC Q59505;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Prepiliin.

OS Moraxella bovis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Moraxellaceae; Moraxella.

OX NCBI_TaxId=476;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=TAT849 serogroup E;

RA MEDLINE=94327452; PubMed=8051000;

RT "Characterization of piliin genes from seven serologically defined RT prototype strains of Moraxella bovis."

RL J. Bacteriol. 176:4875-4882(1994).

EMBL: L32972; AAAS3561.1; -; Genomic_DNA.

DR HSP: P02973; 100W.

DR GO: GO:0009289; C:fimbrium; IEA.

DR GO: GO:0015627; C:type II protein secretion system complex; IEA.

DR GO: GO:0008565; F:protein transporter activity; IEA.

DR GO: GO:0007155; P:cell adhesion; IEA.

DR GO: GO:0015628; P:type II protein secretion system; IEA.

DR InterPro: IPR000983; Bac_GSPH.

DR InterPro: IPR002416; Piliin.

DR InterPro: IPR001082; Piliin.

DR Pfam: PF00114; Piliin; 1.

DR PRINTS: PR00813; BCTERIALGSPH.

DR ProDom: PD000666; Piliin; 1.

DR PROSITE: PS00409; PROKAR_NTER_METHYL; 1.

KM Fimbria.

SO SEQUENCE 159 AA; 16534 MW; F40828A90B93089 CRC64;

Query Match 96.5%; Score 111; DB 2; Length 159;

Best Local Similarity 92.0%; Pred. No. 7e-07;

Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAITIGILAAIALPAYQDYI 25

DB 10 IELMTVIAITIGILAAIALPAYQDYI 34

RESULT 12

Q59504 MORBO PRELIMINARY; PRT; 160 AA.

ID Q59504 MORBO PRELIMINARY; PRT; 160 AA.

AC Q59504;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Prepiliin.

OS Moraxella bovis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Moraxellaceae; Moraxella.

OX NCBI_TaxId=476;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=S276R serogroup A;

RA MEDLINE=94327452; PubMed=8051000;

RT "Characterization of piliin genes from seven serologically defined

```

RT  prototype strains of Moraxella bovis.";
RL  J. Bacteriol. 176:4875-4882(1994).
DR  EMBL, U32968; AA53560.1; -; Genomic_DNA.
DR  HSSP; P02973; 100M.
DR  GO; GO:0009289; C:flmbrium; IEA.
DR  GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR  GO; GO:0008565; F:protein transporter activity; IEA.
DR  GO; GO:0007155; P:cell adhesion; IEA.
DR  GO; GO:0015628; P:type II protein secretion system; IEA.
DR  InterPro; IPR009833; Bac_GSPH.
DR  InterPro; IPR002416; Bac_GSPH.
DR  InterPro; IPR001120; Prok_N_methyl_S.
DR  Pfam; PF00114; Piliin; 1.
DR  PRINTS; PR00813; BCTRIALGSPG.
DR  PRINTS; PR00885; BCTRIALGSPH.
DR  ProDom; PD000666; Piliin; 1.
DR  PROSITE; PS00409; PROKAR_NTER_METHYL, 1.
KW  Fimbria.
SQ  SEQUENCE 160 AA; 16548 MW; B39939B7A10BF20B CRC64;

Query Match 96.5%; Score 111; DB 2; Length 160;
Best Local Similarity 92.0%; Pred. No. 7, 1e-07;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 1 IELMIVAIIGILAAIALPAYOEYV 25
Db 10 IELMIVAIIGILAAIALPAYODYI 34

RESULT 13
O8EJP5 SHEON PRELIMINARY; PRT; 138 AA.
ID O8EJP5
AC O8EJP5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Piliin, putative.
GN OrderedLocustNames=SO0417;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Shewanellaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.B., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eilen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Dancherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Hafe D.H., Koonsey J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Imprial M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Utecherback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Neilson K.H., Fraser C.M.;
RA "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015490; AA53500.1; -; Genomic_DNA.
DR HSSP; P02973; 100M.
DR TIGR; SO0417; -.
DR GO; GO:0009289; C:flmbrium; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR PRINTS; PR00885; BCTRIALGSPH.
DR ProDom; PD000666; Piliin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL, 1.

```

```

KW Complete proteome; Fimbria.
SQ SEQUENCE 138 AA; 13894 MW; CD28162F3A89BA9 CRC64;

Query Match 94.8%; Score 109; DB 2; Length 138;
Best Local Similarity 88.0%; Pred. No. 1, 2e-06;
Matches 22; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy 1 IELMIVAIIGILAAIALPAYOEYV 25
Db 16 IELMIVAIIGILAAIALPAYOEYV 40

RESULT 14
O9R4K1 EIKCO PRELIMINARY; PRT; 32 AA.
ID O9R4K1 EIKCO
AC O9R4K1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Type IV piliin (Fragment).
OS Eikenella corrodens.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Eikenella.
OX NCBI_TaxID=539;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=95369934; PubMed=7642307;
RA Hood B.L., Hirschberg R.;
RL Infect. Immun. 63:3693-3696(1995).
DR GO; GO:0009289; C:flmbrium; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
SQ SEQUENCE 32 AA; 3477 MW; D54B8E08D76B5AB CRC64;

Query Match 93.9%; Score 108; DB 2; Length 32;
Best Local Similarity 95.8%; Pred. No. 5, 2e-07;
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 IELMIVAIIGILAAIALPAYOEY 24
Db 4 IELMIVAIIGILAAIALPAYODY 27

RESULT 15
P78000 9NEIS PRELIMINARY; PRT; 144 AA.
ID P78000 9NEIS
AC P78000
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE KdsB (KdsB) (KdsC).
GN Name=kdsB; Synonyms=kdsA, kdsC;
OS Kingella denitrificans.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Kingella.
OX NCBI_TaxID=502;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33394;
RX MEDLINE=97101010; PubMed=8945537;
RA Weir S., Lee L.W., Mairs C.F.;
RT "Identification of four complete type 4 piliin genes in a single
RT Kingella denitrificans genome.";
RL Infect. Immun. 64:4993-4999(1996).
DR EMBL; U68761; AB48593.1; -; Genomic_DNA.
DR EMBL; U68760; AB48592.1; -; Genomic_DNA.
DR HSSP; P02974; 1A2.
DR GO; GO:0009289; C:flmbrium; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR002416; Bac_GSPH.

```


DR InterPro; IPR001082; PiliN.
 DR InterPro; IPR001120; Prok_N_methyl_S.
 DR Pfam; PF00114; PiliN; 1.
 DR PRINTS; PR00885; BCTERIALGSPH.
 DR PRODOM; PD000666; PiliN; 1.
 DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
 KW Fimbria.
 SQ SEQUENCE 144 AA; 14718 MW; 849190D02F92A0AF CRC64;

Query Match 93.9%; Score 108; DB 2; Length 144;
 Best Local Similarity 95.8%; Pred. No. 1.6e-06;
 Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ELMIVIAITIGITAAIAPAYOEY 24
 |||||
 Db 11 ELMIVIAITIGITAAIAPAYODY 34

Search completed: January 23, 2006, 10:40:53
 Job time : 80.6943 secs

The Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:52:43 ; Search time 57.4236 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-13
Perfect score: 115
Sequence: 1 IELMIVAIIGILAIAPAYQEVY 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:*
1: /cgn2_6/pcodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/pcodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/pcodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/pcodata/1/pubppaa/US10_PUBCOMB.pep:*
5: /cgn2_6/pcodata/1/pubppaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/pcodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	25	4	US-10-665-184-13
2	115	100.0	25	5	US-10-942-300-13
3	115	100.0	25	5	US-10-501-838a-13
4	112	97.4	25	4	US-10-665-184-14
5	112	97.4	25	5	US-10-942-300-14
6	112	97.4	25	5	US-10-501-838a-14
7	107	93.0	139	3	US-09-848-616-140
8	107	93.0	139	4	US-10-289-454-2
9	107	93.0	139	4	US-10-050-902-140
10	107	93.0	139	4	US-10-050-898-140
11	107	93.0	139	4	US-10-346-190-2
12	107	93.0	139	4	US-10-289-456-2
13	106	92.2	167	4	US-10-267-682-109
14	106	92.2	167	4	US-10-267-748-109
15	106	92.2	170	5	US-10-472-260-130
16	106	92.2	179	5	US-10-497-846-4
17	102	88.7	149	4	US-10-137-032-96
18	101	87.8	26	3	US-09-086-118-29
19	101	87.8	26	3	US-09-810-997-6
20	101	87.8	26	3	US-09-766-348-9
21	101	87.8	26	4	US-10-034-158-9
22	101	87.8	26	4	US-10-174-209-36
23	101	87.8	26	4	US-10-347-806-29
24	101	87.8	26	4	US-10-441-736-17
25	100	87.0	154	4	US-10-267-682-108
26	100	87.0	154	4	US-10-267-748-108
27	91	79.1	149	6	US-11-019-005-2

28	91	79.1	149	6	US-11-019-005-26	Sequence 26, Appl
29	91	79.1	149	6	US-11-019-005-28	Sequence 28, Appl
30	91	79.1	149	6	US-11-019-005-30	Sequence 30, Appl
31	91	79.1	149	6	US-11-019-005-32	Sequence 32, Appl
32	91	79.1	149	6	US-11-019-005-34	Sequence 34, Appl
33	91	79.1	149	6	US-11-019-005-38	Sequence 38, Appl
34	91	79.1	149	6	US-11-019-005-40	Sequence 40, Appl
35	91	79.1	149	6	US-11-019-005-42	Sequence 42, Appl
36	91	79.1	149	6	US-11-019-005-44	Sequence 44, Appl
37	87	75.7	149	6	US-11-019-005-36	Sequence 36, Appl
38	87	75.7	292	3	US-09-975-719-419	Sequence 419, App
39	86	74.8	145	4	US-10-282-122a-56078	Sequence 56078, A
40	60	52.2	213	3	US-09-975-719-383	Sequence 383, App
41	57	49.6	288	5	US-10-450-763-55321	Sequence 55321, A
42	54.5	47.4	289	4	US-10-437-963-195929	Sequence 195929,
43	54	47.0	298	4	US-10-437-963-202821	Sequence 202821,
44	51	44.3	1011	4	US-10-276-774-2143	Sequence 2143, Ap
45	50	43.5	107	4	US-10-289-762-982	Sequence 982, App

ALIGNMENTS

RESULT 1
US-10-665-184-13
; Sequence 13, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Klugeella dentrificans
US-10-665-184-13
Query Match 100.0%; Score 115; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 IELMIVAIIGILAIAPAYQEVY 25
Db 1 IELMIVAIIGILAIAPAYQEVY 25
RESULT 2
US-10-942-300-13
; Sequence 13, Application US/10942300
; Publication No. US2005016103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 13
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Kingella denitrificans
US-10-942-300-13

Query Match 100.0%; Score 115; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMTVIAIIGILAAIALPAYQEVY 25
Db 1 IELMTVIAIIGILAAIALPAYQEVY 25

RESULT 3
US-10-501-838A-13

;; Sequence 13, Application US/10501838A
;; Publication No. US20050215478A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel A.

;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501 NATL
;; CURRENT APPLICATION NUMBER: US/10/501,838A
;; PRIOR FILING DATE: 2004-07-19
;; PRIOR APPLICATION NUMBER: PCT/IB03/00968
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: US 60/355,396
;; NUMBER OF SEQ ID NOS: 72
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 13
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Kingella denitrificans
US-10-501-838A-13

Query Match 100.0%; Score 115; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMTVIAIIGILAAIALPAYQEVY 25
Db 1 IELMTVIAIIGILAAIALPAYQEVY 25

RESULT 4
US-10-665-184-14

;; Sequence 14, Application US/10665184
;; Publication No. US20040146549A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel
;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501CIP
;; CURRENT APPLICATION NUMBER: US/10/665,184
;; PRIOR FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: PCT/IB03/00968
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: 60/355,396
;; PRIOR FILING DATE: 2002-02-07
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 14
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Eikenella corrodens

US-10-665-184-14

Query Match 97.4%; Score 112; DB 4; Length 25;
Best Local Similarity 96.0%; Pred. No. 3.2e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMTVIAIIGILAAIALPAYQEVY 25
Db 1 IELMTVIAIIGILAAIALPAYQDYV 25

RESULT 5
US-10-942-300-14

;; Sequence 14, Application US/10942300
;; Publication No. US20050136103A1
;; GENERAL INFORMATION:
;; APPLICANT: Cohen, Elina
;; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-503
;; CURRENT APPLICATION NUMBER: US/10/942,300
;; PRIOR FILING DATE: 2004-09-16
;; PRIOR APPLICATION NUMBER: 10/665,184
;; PRIOR FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: 10/664,989
;; PRIOR FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: 60/503,615
;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 14
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Eikenella corrodens
US-10-942-300-14

Query Match 97.4%; Score 112; DB 5; Length 25;
Best Local Similarity 96.0%; Pred. No. 3.2e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMTVIAIIGILAAIALPAYQEVY 25
Db 1 IELMTVIAIIGILAAIALPAYQDYV 25

RESULT 6
US-10-501-838A-14

;; Sequence 14, Application US/10501838A
;; Publication No. US20050215478A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel A.
;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501 NATL
;; CURRENT APPLICATION NUMBER: US/10/501,838A
;; PRIOR FILING DATE: 2004-07-19
;; PRIOR APPLICATION NUMBER: PCT/IB03/00968
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: 60/355,396
;; PRIOR FILING DATE: 2002-02-07
;; NUMBER OF SEQ ID NOS: 72
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 14
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Eikenella corrodens
US-10-501-838A-14

Query Match 97.4%; Score 112; DB 5; Length 25;
Best Local Similarity 96.0%; Pred. No. 3.2e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Oy      1 1ELMIVIAIGILAAIALPAYQRYV 25
         |||||:|||||
Db      1 1ELMIVIAIGILAAIALPAYQDYV 25
```

```

RESULT 7
US-09-848-616-140
Sequence 140, Application US/09848616
Publication No. US20030054010A1
GENERAL INFORMATION:
APPLICANT: Sebbel, Peter
APPLICANT: Dunant, Nicolas
APPLICANT: Bachmann, Martin
APPLICANT: Tissot, Alain
APPLICANT: Lechner, Franziska
TITLE OF INVENTION: Molecular Antigen Array
FILE REFERENCE: 1700.0180002
CURRENT APPLICATION NUMBER: US/09/848,616
CURRENT FILING DATE: 2001-05-05
NUMBER OF SEQ ID NOS: 186
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 140
LENGTH: 139
TYPE: PRT
ORGANISM: Pseudomonas stutzeri
US-09-848-616-140

```

Query Match	93.0%	Score 107;	DB 3;	Length 139;
Best Local Similarity	91.7%	Pred. No. 1.le-07;		
Matches	22;	Conservative	0;	Indels 0;
		Mismatches	0;	Gaps 0;

```

Qy      1 TELMIVAIIGILAAIALPAYQEY 24
        |||||:|||||||:|
Db      12 TELMIVAIIGILAAIALPAYQDY 35

```

```

      RESULT 8
      US-10-289-454-2
      / Sequence 2, Application US/10289454
      / Publication No. US20030157479A1
      / GENERAL INFORMATION:
      / APPLICANT: Bachmann, Martin
      / APPLICANT: Jemling, Gary
      / APPLICANT: Sonderegger, Ino
      / TITLE OF INVENTION: Antigen Arrays for Treatments of Allergic Eosinophilic Diseases
      / FILE REFERENCE: 1700.0360001
      / CURRENT APPLICATION NUMBER: US 10/289,454
      / CURRENT FILING DATE: 2003-02-10
      / PRIOR APPLICATION NUMBER: US 60/396,636
      / PRIOR FILING DATE: 2002-07-19
      / PRIOR APPLICATION NUMBER: PCT/IB02/00166
      / PRIOR FILING DATE: 2002-01-21
      / PRIOR APPLICATION NUMBER: US 10/050,902
      / PRIOR FILING DATE: 2002-01-18
      / PRIOR APPLICATION NUMBER: US 60/331,045
      / PRIOR FILING DATE: 2001-11-07
      / NUMBER OF SEQ ID NOS: 386
      / SOFTWARE: PatentIn version 3.2
      / SEQ ID NO 2
      / LENGTH: 139
      / TYPE: prt
      / ORGANISM: Pseudomonas stutzeri
      US-10-289-454-2

```

Query Match	93.0%	Score 107;	DB 4;	Length 139;
Best Local Similarity	91.7%	Pred. No. 1.1e-07;		
Matches	22;	Conservative	0;	Indels 0;
		Mismatches	0;	Gaps 0;

```
QY      1 IELMIVAIIGILAAIALPAYOEY 24
        |||||:|||||:|||||:|
DB      12 IELMIVAIIGILAAIALPAYQDY 35
```

```

RESULT 9
US-10-050-902-140
; Sequence 140, Application US/10050902
; Publication NO. US20030175290A1
; GENERAL INFORMATION:
; APPLICANT: Renner, Wolfgang A.
; APPLICANT: Bachmann, Martin
; APPLICANT: Tisbet, Alain
; APPLICANT: Maurer, Patrick
; APPLICANT: Lechner, Franziska
; APPLICANT: Seibel, Peter
; APPLICANT: Plosek, Christine
; TITLE OF INVENTION: Molecular Antigen Array
; FILE REFERENCE: 1700.0190004
; CURRENT APPLICATION NUMBER: US/10/050,902
; PRIORITY FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/262,379
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/288,549
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/326,998
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/331,045
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 140
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Pseudomonas stutzeri
US-10-050-902-140

```

Query Match	93.0%	Score 107;	DB 4;	Length 139;
Best Local Similarity	-91.7%;	Pred. NO. 1.1e-07;		
Matches	22;	Conservative	2;	Mismatches 0;
			Indels	0;
			Gaps	0;

```
OY      1 IELMIVAIIGILAIALPAYOEY 24
        |||||:|||||:|||||:|
Db      12 IELMIVAIIGILAIALPAYODY 35
```

```

RESULT 10
US-10-050-898-140
/ Sequence 140, Application US/10050898
/ Publication No. US2003017571A1
/ GENERAL INFORMATION:
/ APPLICANT: Renner, Wolfgang A.
/ APPLICANT: Bachmann, Martin
/ APPLICANT: Tjssoc, Alain
/ APPLICANT: Mauret, Patrick
/ APPLICANT: Lechner, Franziska
/ APPLICANT: Sebbel, Peter
/ APPLICANT: Plossel, Christine
/ APPLICANT: Ortman, Rainer
/ APPLICANT: Iwond, Rainer
/ APPLICANT: Staufenbiel, Matthias
/ APPLICANT: Frey, Peter
/ TITLE OF INVENTION: Molecular Antigen Array
/ FILE REFERENCE: 1700-0190005
/ CURRENT APPLICATION NUMBER: US/10/050-898
/ PRIOR FILING DATE: 2002-01-18
/ PRIOR APPLICATION NUMBER: US 60/262,379
/ PRIOR FILING DATE: 2001-01-19
/ PRIOR APPLICATION NUMBER: US 60/288,549
/ PRIOR FILING DATE: 2001-05-04
/ PRIOR APPLICATION NUMBER: US 60/326,998
/ PRIOR FILING DATE: 2001-10-05
/ PRIOR APPLICATION NUMBER: US 60/331,045
/ PRIOR FILING DATE: 2001-11-07
/ NUMBER OF SEQ ID NOS: 350
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 140

```

LENGTH: 139
TYPE: PRT
ORGANISM: Pseudomonas stutzeri
US-10-050-898-140

Query Match 93.0%; Score 107; DB 4; Length 139;
Best Local Similarity 91.7%; Pred. No. 1.1e-07;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAIAPAYOEY 24
Db 12 IELMIVAIIGILAIAPAYODY 35

RESULT 11
US-10-346-190-2
Sequence 2, Application US/10346190
Publication No. US20030219459A1
GENERAL INFORMATION:
APPLICANT: Bachmann, Martin
APPLICANT: Maurer, Patrick
APPLICANT: Pelliccioli, Erica
TITLE OF INVENTION: Prion Protein Carrier-Conjugates
FILE REFERENCE: 1700.0290003
CURRENT APPLICATION NUMBER: US/10/346,190
CURRENT FILING DATE: 2003-01-17
PRIOR APPLICATION NUMBER: 60/396,590
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/393,725
PRIOR FILING DATE: 2002-07-08
PRIOR APPLICATION NUMBER: 60/389,898
PRIOR FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: PCT/IB02/00166
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: 10/050,902
PRIOR FILING DATE: 2002-01-18
NUMBER OF SEQ ID NOS: 164
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 139
TYPE: PRT
ORGANISM: Pseudomonas stutzeri
US-10-346-190-2

Query Match 93.0%; Score 107; DB 4; Length 139;
Best Local Similarity 91.7%; Pred. No. 1.1e-07;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAIAPAYOEY 24
Db 12 IELMIVAIIGILAIAPAYODY 35

RESULT 12
US-10-289-456-2
Sequence 2, Application US/10289456
Publication No. US20040033211A1
GENERAL INFORMATION:
APPLICANT: Bachmann, Martin
APPLICANT: Maurer, Patrick
APPLICANT: Spohn, Gunther
TITLE OF INVENTION: Antigen Arrays for Treatment of Bone Disease
FILE REFERENCE: 1700.0330001
CURRENT APPLICATION NUMBER: US/10/289,456
CURRENT FILING DATE: 2002-11-07
PRIOR APPLICATION NUMBER: PCT/IB02/00166
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: US 10/050,902
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/396,635
PRIOR FILING DATE: 2002-07-19
PRIOR APPLICATION NUMBER: US 60/331,045

PRIOR FILING DATE: 2001-11-07
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 139
TYPE: PRT
ORGANISM: Pseudomonas stutzeri
US-10-289-456-2

Query Match 93.0%; Score 107; DB 4; Length 139;
Best Local Similarity 91.7%; Pred. No. 1.1e-07;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAIAPAYOEY 24
Db 12 IELMIVAIIGILAIAPAYODY 35

RESULT 13
US-10-267-682-109
Sequence 109, Application US/10267682
Publication No. US2004003235A1
GENERAL INFORMATION:
APPLICANT: Bolognesi, Daniel P.
Matchewes, Thomas J.
Wild, Carl T.
Barney, Shawn O.
Lambert, Dennis M.
Petteaway, Stephen R.
Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TRANSMISSION
NUMBER OF SEQUENCES: 239
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/267,682
FILING DATE: 08-Oct-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/484,223A
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Corlucci, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8664
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 109:
US-10-267-682-109
Query Match 92.2%; Score 106; DB 4; Length 167;
Best Local Similarity 87.5%; Pred. No. 1.9e-07;

Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIGIILALPAYOEY 24
| | | | | : | | | | : | | | | : |
DB 11 IELMTVIAIGIILALPAYODY 34

RESULT 14

US-10-267-748-109
Sequence 109, Application US/10267748
Publication No. US20040052820A1

GENERAL INFORMATION:

APPLICANT: Bolognesi, Dani P.
Matchews, Thomas J.

Wild, Carl T.

Barney, Shawn O.

Lambert, Dennis M.

Petteaway, Stephen R.

Langlois, Alphonse J.

TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV

TRANSMISSION

NUMBER OF SEQUENCES: 239

CORRESPONDENCE ADDRESSES:

ADDRESSER: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/267,748

FILING DATE: 08-Oct-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/484,223A

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7872-029

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 109:

SEQUENCE CHARACTERISTICS:

LENGTH: 167 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: unknown

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 109:

US-10-267-748-109

Query Match 92.2%; Score 106; DB 4; Length 167;

Best Local Similarity 87.5%; Pred. No. 1.9e-07;

Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIGIILALPAYOEY 24
| | | | | : | | | | : | | | | : |

DB 11 IELMTVIAIGIILALPAYODY 34

RESULT 15

US-10-472-260-130

Sequence 130, Application US/10472260

Publication No. US20040265328A1

GENERAL INFORMATION:

APPLICANT: MICROBIOLOGICAL RESEARCH AUTHORITY

IMPERIAL COLLEGE INNOVATIONS LIMITED

TITLE OF INVENTION: IMMUNOGENIC COMMENSAL NEISSERIA SEQUENCES

FILE REFERENCE: GMS/DJC/23480

CURRENT APPLICATION NUMBER: US/10/472,260

CURRENT FILING DATE: 2003-09-22

NUMBER OF SEQ ID NOS: 199

SOFTWARE: Patentin version 3.1

SEQ ID NO 130

LENGTH: 170

TYPE: prt

ORGANISM: Neisseria meningitidis (group B)

US-10-472-260-130

Query Match 92.2%; Score 106; DB 5; Length 170;

Best Local Similarity 87.5%; Pred. No. 1.9e-07;

Matches 21; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIGIILALPAYOEY 24
| | | | | : | | | | : | | | | : |

DB 11 IELMTVIAIGIILALPAYODY 34

Search completed: January 23, 2006, 12:08:59

Job time : 58.4236 secs

Handwritten text, possibly a signature or date, oriented diagonally.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:32:56 ; Search time 13.6463 Seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-14

Sequence: 1 IELMTVIAIIGTAAIAPAYQDYV 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	99.1	154	2 A44809	fibribial protein -
2	115	99.1	155	2 J10071	alpha-pilin - Mora
3	115	99.1	156	2 A41490	beta fibribial prot
4	115	99.1	157	2 A24434	prepilin - Moraxel
5	115	99.1	158	2 A55851	fibribial protein Q
6	115	99.1	164	2 A42460	pilin pilB precurs
7	110	94.8	159	2 S22064	type 4 pilin - Blk
8	110	94.8	159	2 A47699	fibribial protein c
9	110	94.8	160	2 S40063	fibribial protein c
10	110	94.8	160	2 S74317	fibribial protein c
11	110	94.8	160	2 S74318	fibribial protein c
12	110	94.8	163	2 S40060	fibribial protein c
13	110	94.8	163	2 S40061	fibribial protein c
14	110	94.8	163	2 S40062	fibribial protein c
15	110	94.8	166	1 YQNRG	fibribial protein c
16	110	94.8	166	2 S55496	fibribial protein p
17	110	94.8	167	1 S03091	fibribial protein c
18	110	94.8	167	2 A29611	fibribial protein p
19	110	94.8	168	2 S57418	fibribial protein M
20	110	94.8	168	2 S22102	fibribial protein M
21	110	94.8	168	2 S57416	fibribial protein M
22	110	94.8	168	2 S22103	fibribial protein M
23	110	94.8	169	2 S22104	fibribial protein M
24	110	94.8	170	2 D82021	fibribial protein p
25	110	94.8	170	2 F81246	pilin pilE NMB0018
26	110	94.8	180	2 S1537	fibribial protein M
27	110	94.8	214	2 S15326	fibribial protein M
28	110	94.8	157	2 A31105	fibribial protein p
29	106	91.4	156	2 S15638	pilin (variant NGF

30	106	91.4	157	2 S15637	pilin (variant NGF
31	106	91.4	157	2 S15642	pilin (variant NGF
32	106	91.4	159	2 S15640	pilin (variant NGF
33	106	91.4	159	2 S15639	pilin (variant NGF
34	106	91.4	164	2 S15641	pilin (variant NGF
35	106	91.4	168	2 S22107	pilin pilE precurs
36	105	90.5	85	2 B24886	P(-)rp(-) fibribial
37	105	90.5	136	2 S52692	fibribial protein - Xanth
38	105	90.5	166	2 D24886	P(-)rp(+) fibribial
39	104	89.7	154	2 B31105	fibribial protein p
40	104	89.7	154	2 S04440	fibribial protein -
41	103	88.8	149	2 A25023	type 4 fibribial pr
42	103	88.8	150	1 YQSPSA	fibribial protein p
43	103	88.8	150	2 B24603	fibribial protein -
44	103	88.8	150	2 A43504	pilin precursor -
45	103	88.8	153	2 B60891	pilin - Dicheiobac

ALIGNMENTS

RESULT 1
A44809
fibribial protein - Moraxella nonliquefaciens
M:Alternate names: 198 antigen; pilin
C:Species: Moraxella nonliquefaciens
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A44809; A05122
R:Toujunt, T.; Marrs, C.F.; Rozsa, F.; Boyre, K.
J. Gen. Microbiol. 137, 2483-2490, 1991
A:Title: The type 4 pilin of Moraxella nonliquefaciens exhibits unique similarities wtl
A:Reference number: A44809; PMID:92121902; PMID:1770363
A:Accession: A44809
A:Molecule type: DNA
A:Residues: 1-154 <TON>
A:Cross-references: UNIPROT:P09829; UNIPARC:UPI000012AAA7
A:Experimental source: strain NCCT 7784
A:Note: Sequence extracted from NCBI backbone (NCBIN:77471, NCBIP:77472)
R:Proholm, L.O.; Sletten, K.
FEBS Lett. 73, 29-32, 1977
A:Title: Purification and N-terminal sequence of a fibribial protein from Moraxella nonl
A:Reference number: A05122; PMID:77116125; PMID:838045
A:Accession: A05122
A:Molecule type: protein
A:Residues: 7-55 <FRO>
A:Cross-references: UNIPARC:UPI00001781F3
C:Genetics:
A:Gene: tfrp
C:Superfamily: gonococcal fibribial protein
C:Keywords: methylated amino end; surface antigen
F:7/Modified site: methylated amino end (Phe) (in mature form) #status experimental

Query Match 99.1%; Score 115; DB 2; Length 154;
Best Local Similarity 96.0%; Pred. No. 2e+08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTAAIAPAYQDYV 25
DB 10 IELMTVIAIIGTAAIAPAYQDYI 34

RESULT 2
J10071
alpha-pilin - Moraxella bovis (strain Bpp63)
C:Species: Moraxella bovis
A:Note: host Bos primigenius taurus (cattle)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: J10071
R:Ruehl, W.W.; Marrs, C.F.; Fernandez, R.; Falkow, S.; Schoolnik, G.K.
J. Exp. Med. 168, 983-1002, 1988
A:Title: Purification, characterization, and pathogenicity of Moraxella bovis pil.
A:Reference number: J10071; PMID:89010522; PMID:2902184
A:Accession: J10071

A:Molecule type: protein
A:Residues: 1-155 <RUB>
A:Crosss-references: UNIPROT:P20657; UNIPARC:UPI00001781F5
A>Note: about 50% of the amino-terminus is N-methylated; the remaining 50% of the molecu
A:Superfamily: gonococcal fimbril protein
C:Keywords: methylated amino end
F:1-155/Product: alpha-pilin 1 #status experimental <MAT1>
F:2-155/Product: alpha-pilin 2 #status experimental <MAT2>
F:1/Modified site: methylated amino end (Phe) #status experimental
F:13-150/Disulfide bonds: #status experimental

Query Match 99.1%; Score 115; DB 2; Length 155;
Best Local Similarity 96.0%; Pred. No. 2e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IELMIVAIIGILAAIAPAYQDYV 25
Db 4 IELMIVAIIGILAAIAPAYQDYI 28

RESULT 3
A:1490
A:pin precursor - Moraxella bovis
C:Species: Moraxella bovis
C>Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 09-Jul-2004
C:Accession: A41490
R:Elleman, T.C.; Hoyle, P.A.; Lepper, A.W.D.
J. Infect. Immun. 58, 1678-1684, 1990
A:Title: Characterization of the pilin gene of Moraxella bovis Dalton 2d and expression
A:Reference number: A41490; MUID:90256243; PMID:1971258
A:Accession: A41490
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <ELI>
A:Crosss-references: UNIPROT:Q59507; UNIPARC:UPI0000B2378; GB:M92155; GB:M32665; NID:914
C:Superfamily: gonococcal fimbril protein
C:Keywords: methylated amino end
F:7/Modified site: methylated amino end (Phe) (in mature form) #status predicted

Query Match 99.1%; Score 115; DB 2; Length 156;
Best Local Similarity 96.0%; Pred. No. 2e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IELMIVAIIGILAAIAPAYQDYV 25
Db 10 IELMIVAIIGILAAIAPAYQDYI 34

RESULT 4
A:24434
A:beta fimbril protein precursor - Moraxella bovis
N:Alternate names: beta pilin
C:Species: Moraxella bovis
C>Date: 08-Aug-1987 #sequence_revision 08-Aug-1987 #text_change 09-Jul-2004
C:Accession: A24434; J00072
R:Marra, C.F.; Schoolnik, G.; Komey, J.M.; Hardy, J.; Rothbard, J.; Falkow, S.
J. Bacteriol. 163, 132-139, 1985
A:Reference number: A24434; MUID:85234350; PMID:2861194
A:Accession: A24434
A:Molecule type: DNA
A:Residues: 1-157 <MAR>
A:Crosss-references: UNIPROT:P07640; UNIPARC:UPI000012AAC7; GB:M11435; NID:9149760; PIDN:
A:Experimental source: strain EPP63
R:Rush, W.W.; Marra, C.F.; Fernandez, R.; Falkow, S.; Schoolnik, G.K.
J. Exp. Med. 168, 983-1002, 1988
A:Title: Purification, characterization, and pathogenicity of Moraxella bovis pilin.
A:Reference number: J00071; MUID:89010522; PMID:2902184
A:Accession: J00072
A:Molecule type: protein
A:Residues: 7-86, 93-122 <RUB>
A:Crosss-references: UNIPARC:UPI00001781F2
A>Note: about 50% of the amino-terminus is N-methylated; the remaining 50% of the molecu

C:Superfamily: gonococcal fimbril protein
C:Keywords: methylated amino end; surface antigen
F:1-6/Domain: propeptide #status predicted <PRO>
F:7-157/Product: beta fimbril protein 1 #status experimental <MAT>
F:8-157/Product: beta fimbril protein 2 #status experimental <MAT2>
F:7/Modified site: methylated amino end (Phe) (in mature form) #status experimental
F:136-155/Disulfide bonds: #status experimental

Query Match 99.1%; Score 115; DB 2; Length 157;
Best Local Similarity 96.0%; Pred. No. 2e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IELMIVAIIGILAAIAPAYQDYV 25
Db 10 IELMIVAIIGILAAIAPAYQDYI 34

RESULT 5
A:55851
A:prepilin - Moraxella bovis
C:Species: Moraxella bovis
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: A55851
R:Atwell, J.L.; Tennent, J.M.; Lepper, A.W.; Elleman, T.C.
J. Bacteriol. 176, 4875-4882, 1994
A:Title: Characterization of pilin genes from seven serologically defined prototype stra
A:Reference number: A55851; MUID:94327452; PMID:8051000
A:Accession: A55851
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-158 <RES>
A:Crosss-references: UNIPROT:Q59501; UNIPARC:UPI0000BACSD; GB:L32965; NID:9486303; PIDN:
C:Superfamily: gonococcal fimbril protein
C:Keywords: methylated amino end
F:7/Modified site: methylated amino end (Phe) (in mature form) #status predicted

Query Match 99.1%; Score 115; DB 2; Length 158;
Best Local Similarity 96.0%; Pred. No. 2e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IELMIVAIIGILAAIAPAYQDYV 25
Db 10 IELMIVAIIGILAAIAPAYQDYI 34

RESULT 6
A:42460
A:fimbril protein Q - Moraxella lacunata (ATCC 17956) plasmid pMX1
C:Species: Moraxella lacunata
C>Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 29-Jan-1999
C:Accession: A42460
R:Roza, F.W.; Marra, C.F.
J. Bacteriol. 173, 4000-4006, 1991
A:Title: Interesting sequence differences between the pilin gene inversion regions of Mo
A:Reference number: A42460; MUID:91286182; PMID:2061282
A:Accession: A42460
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-164 <ROZ>
A:Crosss-references: UNIPARC:UPI00001781F4; GB:M59711
A:Genetics:
A:Genome: Plasmid
C:Superfamily: gonococcal fimbril protein
C:Keywords: methylated amino end
F:7/Modified site: methylated amino end (Phe) (in mature form) #status predicted

Query Match 99.1%; Score 115; DB 2; Length 164;
Best Local Similarity 96.0%; Pred. No. 2.1e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IELMIVAIIGILAAIAPAYQDYV 25
Db 10 IELMIVAIIGILAAIAPAYQDYI 34

RESULT 7

S22064
piliin piliE precursor - Neisseria gonorrhoeae (fragment)
C/Species: Neisseria gonorrhoeae
A/Variety: strain MS11
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C/Accession: S22064
R/Haas, R.
Submitted to the EMBL Data Library, June 1992
A/Reference number: S22064
A/Accession: S22064
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-159 <HAAS>
A/Cross-references: UNIPROT:Q51321; UNIPARC:UPI00000BD9D7; EMBL:X66620; NID:G45254; PIDD
C/Superfamily: gonococcal fimbrial protein

Query Match 94.8%; Score 110; DB 2; Length 159;

Best Local Similarity 91.7%; Pred. No. 9.6e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITIGILAAALPAYQDY 24

Db 4 IELMIVIAITIGILAAALPAYQDY 27

RESULT 8

A47699
type 4 piliin - Eikenella corrodens
N/Alternate names: N-methylphenylalanine-type piliin
C/Species: Eikenella corrodens
C/Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A47699; S23844
R/Rao, V.K.; Progulski-Fox, A.
J. Gen. Microbiol. 139, 651-660, 1993
A/Title: Cloning and sequencing of two type 4 (N-methylphenylalanine) piliin genes from E.
A/Reference number: A47699; MUID:93232782; PMID:8473871
A/Accession: A47699
A/Molecule type: DNA
A/Residues: 1-159 <RAO>
A/Cross-references: UNIPROT:P35645; UNIPARC:UPI0000129B89; EMBL:Z12609; NID:G41329; PIDD
A/Experimental source: ATCC 23834
A/Note: sequence extracted from NCBI backbone (NCBIP:130001)
C/Genetics:
A/Gene: ecpA
C/Superfamily: gonococcal fimbrial protein
C/Keywords: methylated amino end
F/8/Modified site: methylated amino end (Phe) (1n mature form) #status predicted

Query Match 94.8%; Score 110; DB 2; Length 159;

Best Local Similarity 92.0%; Pred. No. 9.6e-08;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IELMIVIAITIGILAAALPAYQDY 25

Db 11 IELMIVIAITIGILAAALPAYQDY 35

RESULT 9

S40063
fimbrial protein class I - Neisseria meningitidis (fragment)
N/Alternate names: piliin
C/Species: Neisseria meningitidis
A/Variety: strain C311 variant 28; strain C311 variant 3
C/Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S40063; S74316
R/Virtj, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A/Title: Pilius-facilitated adherence of Neisseria meningitidis to human epithelial and e
ence and the glycosylation status of piliin.
A/Reference number: S40060; MUID:95020562; PMID:7934852

A/Accession: S40063
A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-160 <VIR>

A/Cross-references: UNIPROT:Q57135; UNIPARC:UPI00000BA7AA; EMBL:L22678; NID:G349523; PIDD

A/Experimental source: strain C311 variant 28

A/Accession: S74316

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-160 <VIR>

A/Cross-references: UNIPARC:UPI00000BA7AA; EMBL:L22636; NID:G349013; PIDD:AAA67334.1; PIDD

A/Experimental source: strain C311 variant 3

C/Genetics:

A/Gene: piliB

C/Superfamily: gonococcal fimbrial protein

C/Keywords: fimbria

Query Match 94.8%; Score 110; DB 2; Length 160;

Best Local Similarity 91.7%; Pred. No. 9.6e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITIGILAAALPAYQDY 24

Db 4 IELMIVIAITIGILAAALPAYQDY 27

RESULT 10

S74317
fimbrial protein class I - Neisseria meningitidis (strain C311 variant 7) (fragment)
N/Alternate names: piliin
C/Species: Neisseria meningitidis
A/Variety: strain C311 variant 7
C/Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S74317
R/Virtj, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A/Title: Pilius-facilitated adherence of Neisseria meningitidis to human epithelial and
ence and the glycosylation status of piliin.
A/Reference number: S40060; MUID:95020562; PMID:7934852
A/Accession: S74317
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-160 <VIR>
A/Cross-references: UNIPROT:Q51234; UNIPARC:UPI00000A8DF1; EMBL:L22676; NID:G349519; PIDD
C/Genetics:
A/Gene: piliB
C/Superfamily: gonococcal fimbrial protein
C/Keywords: fimbria

Query Match 94.8%; Score 110; DB 2; Length 160;

Best Local Similarity 91.7%; Pred. No. 9.6e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITIGILAAALPAYQDY 24

Db 4 IELMIVIAITIGILAAALPAYQDY 27

RESULT 11

S74318
fimbrial protein class I - Neisseria meningitidis (strain C311 variant 16) (fragment)
N/Alternate names: piliin
C/Species: Neisseria meningitidis
A/Variety: strain C311 variant 16
C/Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S74318
R/Virtj, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A/Title: Pilius-facilitated adherence of Neisseria meningitidis to human epithelial and
ence and the glycosylation status of piliin.
A/Reference number: S40060; MUID:95020562; PMID:7934852
A/Accession: S74318
A/Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 1-160 <VIR>
A:Cross-references: UNIPROT:Q51235; UNIPARC:UPI00000BBE20; EMBL:L22677; NID:g349521; PID
C:Genetics:
A:Gene: p11E
C:Superfamily: gonococcal fimbrial protein
C:Keywords: fimbria

Query Match 94.8%; Score 110; DB 2; Length 160;
Best Local Similarity 91.7%; Pred. No. 9.8e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGTIAAALPAYQDY 24
|||||:|||||:|||||
Db 4 IELMIVIAIGTIAAVALPAYQDY 27

RESULT 12
S40060 fimbrial protein class I - Neisseria meningitidis (strain MCS8 variant 5) (fragment)
N:Alternate names: p11n
C:Species: Neisseria meningitidis
A:Variety: strain MCS8 variant 5
C:Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S40060
R:Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A:Title: Pilus-facilitated adherence of Neisseria meningitidis to human epithelial and e
ence and the glycosylation status of p11n.
A:Reference number: S40060; MUID:95020562; PMID:7934852
A:Accession: S40060
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-163 <VIR>
A:Cross-references: UNIPROT:Q51237; UNIPARC:UPI00000BEBD3; EMBL:L22681; NID:g349015; PID
C:Genetics:
A:Gene: p11E
C:Superfamily: gonococcal fimbrial protein
C:Keywords: fimbria

Query Match 94.8%; Score 110; DB 2; Length 163;
Best Local Similarity 91.7%; Pred. No. 9.8e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGTIAAALPAYQDY 24
|||||:|||||:|||||
Db 4 IELMIVIAIGTIAAVALPAYQDY 27

RESULT 13
S40061 fimbrial protein class I - Neisseria meningitidis (strain MCS8 variant 6) (fragment)
N:Alternate names: p11n
C:Species: Neisseria meningitidis
A:Variety: strain MCS8 variant 6
C:Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S40061
R:Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A:Title: Pilus-facilitated adherence of Neisseria meningitidis to human epithelial and e
ence and the glycosylation status of p11n.
A:Reference number: S40060; MUID:95020562; PMID:7934852
A:Accession: S40061
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-163 <VIR>
A:Cross-references: UNIPROT:Q51236; UNIPARC:UPI00000AFEBD; EMBL:L22679; NID:g349525; PID
C:Genetics:
A:Gene: p11E
C:Superfamily: gonococcal fimbrial protein
C:Keywords: fimbria

Query Match 94.8%; Score 110; DB 2; Length 163;

Best Local Similarity 91.7%; Pred. No. 9.8e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGTIAAALPAYQDY 24
|||||:|||||:|||||
Db 4 IELMIVIAIGTIAAVALPAYQDY 27

RESULT 14
S40062 fimbrial protein class I - Neisseria meningitidis (strain MCS8 variant 7) (fragment)
N:Alternate names: p11n
C:Species: Neisseria meningitidis
A:Variety: strain MCS8 variant 7
C:Date: 13-Jan-1995 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S40062
R:Virji, M.; Saunders, J.R.; Sims, G.; Makepeace, K.; Maskell, D.; Ferguson, D.J.P.
Mol. Microbiol. 10, 1013-1028, 1993
A:Title: Pilus-facilitated adherence of Neisseria meningitidis to human epithelial and e
ence and the glycosylation status of p11n.
A:Reference number: S40060; MUID:95020562; PMID:7934852
A:Accession: S40062
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-163 <VIR>
A:Cross-references: UNIPROT:Q57441; UNIPARC:UPI00000BDE1C; EMBL:L22683; NID:g349527; PID
C:Genetics:
A:Gene: p11E
C:Superfamily: gonococcal fimbrial protein
C:Keywords: fimbria

Query Match 94.8%; Score 110; DB 2; Length 163;
Best Local Similarity 91.7%; Pred. No. 9.8e-08;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGTIAAALPAYQDY 24
|||||:|||||:|||||
Db 4 IELMIVIAIGTIAAVALPAYQDY 27

RESULT 15
YONHG fimbrial protein class I precursor - Neisseria gonorrhoeae
N:Alternate names: MS11 antigen; p11n
C:Species: Neisseria gonorrhoeae
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004
C:Accession: A94007; A92776; A44149; S13372; C24886; A24886; E24886; A03498
R:Meyer, T.F.; Billiard, E.; Haas, R.; Scorzbach, S.; So, M.
Proc. Natl. Acad. Sci. U.S.A. 81, 6110-6114, 1984
A:Title: Pilus genes of Neisseria gonorrhoeae: chromosomal organization and DNA sequence.
A:Reference number: A94007; MUID:85014915; PMID:6148752
A:Accession: A94007
A:Molecule type: DNA
A:Residues: 1-166 <MEX>
A:Cross-references: UNIPROT:P02974; UNIPARC:UPI000016FBFF; GB:K02078; NID:g150284; PIDN:
A:Experimental source: strain MS11
R:Schoolnik, G.K.; Fernandez, R.; Tai, J.Y.; Rochard, J.; Gotschlich, E.C.
J. Exp. Med. 159, 1351-1370, 1984
A:Title: Gonococcal p11. Primary structure and receptor binding domain.
A:Reference number: A92776; MUID:84187404; PMID:6143785
A:Accession: A92776
A:Molecule type: protein
A:Residues: 8-166 <SCH>
A:Cross-references: UNIPARC:UPI0000141D34
A:Experimental source: strain MS11
A:Note: phosphoserine was detected but not located
R:Johnson, A.B.; Pfeiffer, J.; Normark, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 3204-3208, 1992
A:Title: Neisseria gonorrhoeae PilC expression provides a selective mechanism for struct
A:Reference number: A44149; MUID:92228752; PMID:11348857
A:Accession: A44149
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: DNA

A/Residues: 1-166 <JON1>
 A/Cross-references: UNIPARC:UPI000016FBFP
 A/Experimental source: strain MS1mk[P+]-u
 R.Jonsson, A.B.; Nyberg, G.; Normark, S.
 EMBO J. 10, 477-488, 1991
 A/Title: Phase variation of gonococcal pil1 by frameshift mutation in pilC, a novel gene
 A/Reference number: S13372; NCID:91122056; PMID:1671354
 A/Accession: S13372
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-166 <JON2>
 A/Cross-references: UNIPARC:UPI000016FBFP
 R.Bergstrom, S.; Robbins, K.; Koomen, J.M.; Swanson, J.
 Proc. Natl. Acad. Sci. U.S.A. 83, 3890-3894, 1986
 A/Title: Piliatlon control mechanisms in *Neisseria gonorrhoeae*.
 A/Reference number: A24886; NCID:86233338; PMID:2872674
 A/Accession: C24886
 A/Molecule type: DNA
 A/Residues: 1-24, 'L', 26-166 <BER1>
 A/Cross-references: UNIPARC:UPI00001747DF; GB:M13222; NCID:G150290
 A/Experimental source: P(-)rp(-) revertant P(+) rev
 A/Accession: A24886
 A/Molecule type: DNA
 A/Residues: 1-24, 'L', 26-60, 'N', 62, 'T', 64-68, 'A', 70-73, 'G', 75, 'S', 77, 'T', 79-158, 'L', 160-1
 A/Cross-references: UNIPARC:UPI00001747B0; GB:M13222; NCID:G150290
 A/Experimental source: P(++)
 A/Accession: E24886
 A/Molecule type: DNA
 A/Residues: 1-24, 'L', 26-60, 'N', 62, 'T', 64-68, 'A', 70-85, 'QK', 88-90, 'AK', 93-99, 'A', 101-102,
 A/Cross-references: UNIPARC:UPI00001747B1; GB:M13222; NCID:G150290
 A/Experimental source: P(-)rp(+) revertant P(++)
 A/Function:
 C/Description: surface protein that binds various host cell molecules
 C/Superfamily: gonococcal fimbrial protein
 C/Keywords: fimbria; glycoprotein; methylated amino end; phosphoprotein; surface antigen
 F/1-7/Domain: propeptide #status predicted <PRO>
 F/8-166/Product: fimbrial protein #status predicted <MAT>
 F/31-111/Domain: receptor binding #status predicted <RBC>
 F/8/Modified site: methylated amino end (Phe) (in mature form) #status experimental
 F/101/Binding site: sn-1-glycerolphosphate (Ser) (covalent) #status predicted
 F/128-158/Diulfide bonds: #status experimental

Query Match 94.8%; Score 110; DB 1; Length 166;
 Best local Similarity 91.7%; Pred. No. 9.9e-08;
 Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVIAITGIIAAIALPAYODY 24
 |||||:|||||
 DB 11 IELMIVIAITGIIAAIALPAYODY 34

Search completed: January 23, 2006, 10:43:07
 Job time: 13.6463 secs

The 2000 Bank (over)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:35:50 ; Search time 22.9258 seconds
(without alignments)
90.156 Million cell updates/sec

Title: US-10-501-838A-14
Perfect score: 116
Sequence: 1 IELMIVIAIGILAAIALPAYODY 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents, AA:*
1: /cgn2_6/prodata/1/1aa/5/COMB.pep.*
2: /cgn2_6/prodata/1/1aa/6/COMB.pep.*
3: /cgn2_6/prodata/1/1aa/7/COMB.pep.*
4: /cgn2_6/prodata/1/1aa/8/COMB.pep.*
5: /cgn2_6/prodata/1/1aa/9/COMB.pep.*
6: /cgn2_6/prodata/1/1aa/10/COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	94.8	166	2	US-09-350-841A-1595
2	110	94.8	167	2	US-08-486-099-109
3	110	94.8	167	2	US-08-360-107A-119
4	110	94.8	167	2	US-08-484-223B-109
5	110	94.8	167	2	US-08-919-597-109
6	110	94.8	167	2	US-08-475-668A-109
7	110	94.8	167	2	US-08-485-551A-109
8	110	94.8	167	2	US-08-471-913A-109
9	110	94.8	167	2	US-08-485-264A-109
10	110	94.8	167	2	US-08-474-349A-109
11	110	94.8	167	2	US-08-470-896-109
12	110	94.8	167	2	US-08-485-546A-109
13	110	94.8	167	2	US-08-487-266A-109
14	110	94.8	167	2	US-08-484-741-109
15	109	94.0	174	2	US-09-252-991A-26702
16	108	93.1	159	2	US-09-328-352-6277
17	105	90.5	26	1	US-08-891-254-9
18	105	90.5	26	1	US-08-819-539-9
19	105	90.5	26	1	US-09-030-370A-9
20	105	90.5	26	1	US-08-984-207-9
21	105	90.5	26	2	US-09-013-587-9
22	105	90.5	26	2	US-09-086-118-29
23	105	90.5	26	2	US-09-431-614-17
24	105	90.5	26	2	US-09-412-100-37
25	105	89.7	26	4	PCT-US96-08819-9
26	104	89.7	154	2	US-08-486-099-108
27	104	89.7	154	2	US-08-360-107A-118

28	104	89.7	154	2	US-08-484-223B-108	Sequence 108, App
29	104	89.7	154	2	US-08-919-597-108	Sequence 108, App
30	104	89.7	154	2	US-08-475-668A-108	Sequence 108, App
31	104	89.7	154	2	US-08-485-551A-108	Sequence 108, App
32	104	89.7	154	2	US-08-471-913A-108	Sequence 108, App
33	104	89.7	154	2	US-08-485-264A-108	Sequence 108, App
34	104	89.7	154	2	US-08-474-349A-108	Sequence 108, App
35	104	89.7	154	2	US-08-470-896-108	Sequence 108, App
36	104	89.7	154	2	US-08-485-546A-108	Sequence 108, App
37	104	89.7	154	2	US-08-487-266A-108	Sequence 108, App
38	104	89.7	154	2	US-08-484-741-108	Sequence 108, App
39	90	77.6	162	2	US-09-303-518D-340	Sequence 340, App
40	88	75.9	178	2	US-09-489-039A-12339	Sequence 12339, A
41	86	74.1	162	2	US-09-303-518D-344	Sequence 344, App
42	86	74.1	256	2	US-09-543-681A-4903	Sequence 4903, App
43	85	73.3	163	2	US-09-543-681A-4903	Sequence 419, App
44	84	72.4	292	2	US-09-199-637A-419	Sequence 6111, App
45	82	70.7	161	2	US-09-328-352-6111	

ALIGNMENTS

```
RESULT 1
US-09-350-841A-1595
; Sequence 1595, Application US/09350841A
; Patent No. 6750008
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350,841A
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1595
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-350-841A-1595

Query Match      94.8%; Score 110; DB 2; Length 166;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 IELMIVIAIGILAAIALPAYODY 24
DB      11 IELMIVIAIGILAAIALPAYODY 34

RESULT 2
US-08-486-099-109
; Sequence 109, Application US/08486099
; Patent No. 6013263
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; TITLE OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
```

ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,099
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-031
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-486-099-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGIIAALPAYDY 24
Db 11 IELMIVIAIGIIAALPAYDY 34

RESULT 3
US-08-360-107A-119
Sequence 119, Application US/08360107A
Patent No. 6017536
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petlewy, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,107A
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-013
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 119:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-360-107A-119

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGIIAALPAYDY 24
Db 11 IELMIVIAIGIIAALPAYDY 34

RESULT 4
US-08-484-223B-109
Sequence 109, Application US/08484223B
Patent No. 6020459
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petlewy, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
NUMBER OF SEQUENCES: 245
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,223B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-223B-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITGILAAIALPAYODY 24
11 IELMIVIAITGILAAIALPAYODY 34
Db

RESULT 5

US-08-919-597-109

Sequence 109, Application US/08919597

Patent No. 6054265

GENERAL INFORMATION:

APPLICANT: Bolognesi, Dani P.

APPLICANT: Matthews, Thomas J.

APPLICANT: Wild, Carl T.

APPLICANT: Barney, Shawn O.

APPLICANT: Lambert, Dennis M.

APPLICANT: Petteway, Stephen R.

APPLICANT: Langlois, Alphonse J.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION

TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV

TITLE OF INVENTION: TRANSMISSION

NUMBER OF SEQUENCES: 273

CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/919,597

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/470,896

FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7872-020

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 109:

SEQUENCE CHARACTERISTICS:

LENGTH: 167 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-919-597-109

Query Match 94.8%; Score 110; DB 2; Length 167;

Best local similarity 91.7%; Pred. No. 2e-09;

Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITGILAAIALPAYODY 24
11 IELMIVIAITGILAAIALPAYODY 34
Db

RESULT 6

US-08-475-668A-109

Sequence 109, Application US/08475668A

Patent No. 6060065

GENERAL INFORMATION:

APPLICANT: Barney, Shawn O.

APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/475,668A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7872-026

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 109:

SEQUENCE CHARACTERISTICS:

LENGTH: 167 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-475-668A-109

Query Match 94.8%; Score 110; DB 2; Length 167;

Best local similarity 91.7%; Pred. No. 2e-09;

Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITGILAAIALPAYODY 24
11 IELMIVIAITGILAAIALPAYODY 34
Db

RESULT 7

US-08-485-551A-109

Sequence 109, Application US/0848551A

Patent No. 6068973

GENERAL INFORMATION:

APPLICANT: Bolognesi, Dani P.

APPLICANT: Matthews, Thomas J.

APPLICANT: Wild, Carl T.

APPLICANT: Barney, Shawn O.

APPLICANT: Lambert, Dennis M.

APPLICANT: Petteway, Stephen R.

APPLICANT: Langlois, Alphonse J.

TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE

TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS

TITLE OF INVENTION: TRANSMISSION

NUMBER OF SEQUENCES: 211

CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,551A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-023
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-551A-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAAIPAYQDY 24
Db 11 IELMIVAIIGILAAIPAYQDY 34

RESULT 8
US-08-471-913A-109
Sequence 109, Application US/08471913A
Patent No. 6093794
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Peteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 214
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,913A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-030
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-471-913A-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAAIPAYQDY 24
Db 11 IELMIVAIIGILAAIPAYQDY 34

RESULT 9
US-08-485-264A-109
Sequence 109, Application US/08485264A
Patent No. 6228983
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Peteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 232
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,264A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-021
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-264A-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAAIPAYQDY 24

Db 11 IELMIVIAIGILAAVALPAYODY 34

RESULT 10
US-08-474-349A-109
Sequence 109, Application US/08474349A
Patent No. 6333395
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
VIRUS TRANSMISSION
TITLE OF INVENTION: VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 517
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,349A
FILING DATE: 07-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-024
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-474-349A-109
Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 IELMIVIAIGILAAVALPAYODY 24
DB 11 IELMIVIAIGILAAVALPAYODI 34
RESULT 11
US-08-470-896-109
Sequence 109, Application US/08470896
Patent No. 6479055
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TRANSMISSION
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,896
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-470-896-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 IELMIVIAIGILAAVALPAYODY 24
DB 11 IELMIVIAIGILAAVALPAYODI 34

RESULT 12
US-08-485-546A-109
Sequence 109, Application US/08485546A
Patent No. 6518013
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
TRANSMISSION
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 214
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/485,546A
APPLICATION NUMBER: US/08/485,546A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-485-546A-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGILAAALPAYODY 24
Db 11 IELMIVIAIGILAAALPAYODY 34

RESULT 13
US-08-487-266A-109
Sequence 109, Application US/08487266A
Patent No. 6824783
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
NUMBER OF SEQUENCES: 239
CORRESPONDENCE ADDRESSES:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,266A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-025
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-487-266A-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGILAAALPAYODY 24
Db 11 IELMIVIAIGILAAALPAYODY 34

RESULT 14
US-08-484-741-109
Sequence 109, Application US/08484741
Patent No. 6951717
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TRANSMISSION
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESSES:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,741
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-022
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 109:
US-08-484-741-109

Query Match 94.8%; Score 110; DB 2; Length 167;
Best Local Similarity 91.7%; Pred. No. 2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAIGILAAALPAYODY 24
Db 11 IELMIVIAIGILAAALPAYODY 34

RESULT 15

US-09-252-991A-26702
Sequence 26702, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26702
LENGTH: 174
TYPE: PRF
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26702

Query Match 94.0%; Score 109; DB 2; Length 174;
Best Local Similarity 91.7%; Pred. No. 3e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ISLMTIVATIGTAAATAPAYODY 24
DB 28 ISLMTIVATIGTAAATAPAYODY 51

Search completed: January 23, 2006, 10:46:48
Job time : 22.9258 secs

Copyright © 2000 by
The McGraw-Hill Companies

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:53:39 ; Search time 3.49345 Seconds
(without alignments)
72.521 Million cell updates/sec

Title: US-10-501-838a-14
Perfect score: 116
Sequence: 1 IELMIVAIIGILAAIAPAYQDYV 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA_New.*
1: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubppaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubppaa/US05_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubppaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	94.8	98	6	US-10-467-657-842
2	110	94.8	170	7	US-11-052-554A-238
3	105	90.5	26	6	US-10-972-587-36
4	92	79.3	149	7	US-11-052-554A-82
5	86	74.1	162	6	US-10-467-657-3494
6	78	67.2	129	6	US-10-467-657-6040
7	64	55.2	139	6	US-10-467-657-688
8	57	49.1	236	6	US-10-467-657-3480
9	47	40.5	109	7	US-11-198-847-143
10	44	37.9	223	6	US-10-467-657-6830
11	44	37.9	337	6	US-10-878-556A-143
12	44	37.9	449	6	US-10-131-826A-224
13	44	37.9	519	6	US-10-523-477-8
14	44	37.9	685	7	US-11-089-551A-2
15	43	37.1	755	6	US-10-517-939-330
16	42	36.2	110	7	US-11-198-847-167
17	42	36.2	110	7	US-11-198-847-173
18	42	36.2	247	7	US-11-000-463-242
19	42	36.2	442	6	US-10-467-657-3900
20	42	36.2	645	6	US-10-821-234-1409
21	41	35.3	522	7	US-11-080-991-104
22	41	35.3	852	6	US-10-645-441-15
23	41	35.3	852	7	US-11-050-804-6
24	40	34.5	148	7	US-10-526-716-4
25	40	34.5	205	6	US-10-793-626-240

26	40	34.5	226	5	US-09-978-360A-429	Sequence 429, App
27	40	34.5	316	6	US-10-485-517-370	Sequence 370, App
28	40	34.5	407	6	US-10-821-234-1389	Sequence 1389, App
29	40	34.5	464	6	US-10-689-742-164	Sequence 164, App
30	40	34.5	628	7	US-11-037-243-86	Sequence 86, App
31	39	33.6	91	7	US-11-082-389-352	Sequence 352, App
32	39	33.6	130	6	US-10-995-561-935	Sequence 935, App
33	39	33.6	191	6	US-10-995-561-937	Sequence 937, App
34	39	33.6	191	7	US-11-159-597-11	Sequence 11, App
35	39	33.6	191	7	US-11-219-359-19	Sequence 19, App
36	39	33.6	208	6	US-10-527-500-53	Sequence 53, App
37	39	33.6	426	6	US-10-873-528-33	Sequence 33, App
38	39	33.6	436	6	US-10-513-639-19	Sequence 19, App
39	39	33.6	494	7	US-11-143-980-54	Sequence 54, App
40	39	33.6	525	7	US-11-074-176-146	Sequence 146, App
41	39	33.6	693	6	US-10-467-657-6176	Sequence 6176, App
42	39	33.6	6893	7	US-11-205-109-14	Sequence 14, App
43	38.5	33.2	107	6	US-10-467-657-3318	Sequence 3318, App
44	38	32.8	22	7	US-11-065-943-44	Sequence 44, App
45	38	32.8	211	6	US-10-965-972-4	Sequence 4, App

ALIGNMENTS

RESULT 1
US-10-467-657-842
Sequence 842, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWing9, version 1.04
SEQ ID NO 842
LENGTH: 98
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-842

Query Match 94.8%; Score 110; DB 6; Length 98;
Best local similarity 91.7%; Pred. No. 4,5e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIIGILAAIAPAYQDY 24
DB 11 IELMIVAIIGILAAIAPAYQDY 34

RESULT 2
US-11-052-554A-238
Sequence 238, Application US/11052554A
Publication No. US2005028866A1
GENERAL INFORMATION:
APPLICANT: Sachdeva, et al.
TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
FILE REFERENCE: 30853/40359A
CURRENT APPLICATION NUMBER: US/11/052,554A
CURRENT FILING DATE: 2005-02-07
PRIOR APPLICATION NUMBER: US 60/589,227
PRIOR FILING DATE: 2004-07-20
PRIOR APPLICATION NUMBER: IN 173/DEL/2004
PRIOR FILING DATE: 2004-02-06

```

; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 238
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Neisseria meningitidis 22491
US-11-052-554A-238

Query Match          94.8%; Score 110; DB 7; Length 170;
Best Local Similarity 91.7%; Pred. No. 8.2e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 1 IELMTVIAIGILAIAPAYDY 24
    |||||:||||:||||:|||||
Db 11 IELMTVIAIGILAIAPAYDY 34

RESULT 3
US-10-972-587-36
; Sequence 36, Application US/10972587
; Publication No. US20050246799A1
; GENERAL INFORMATION:
; APPLICANT: Song, Xiaoling
; APPLICANT: Bartola, Pauline Anne
; APPLICANT: Linderoth, Nora Abiella
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 21829/213
; CURRENT APPLICATION NUMBER: US/10/972,587
; PRIOR FILING DATE: 2004-10-25
; PRIOR APPLICATION NUMBER: 60/335,776
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 09/810,997
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 10/174,209
; PRIOR FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Xanthomonas campestris pv. glycines
US-10-972-587-36

Query Match          90.5%; Score 105; DB 6; Length 26;
Best Local Similarity 91.7%; Pred. No. 5.3e-09;
Matches 22; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 IELMTVIAIGILAIAPAYDY 24
    |||||:||||:||||:|||||
Db 3 IELMTVIAIGILAIAPAYDY 26

RESULT 4
US-11-052-554A-82
; Sequence 82, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82
```

```

; LENGTH: 149
; TYPE: PRT
; ORGANISM: Haemophilus influenzae Rd
US-11-052-554A-82

Query Match          79.3%; Score 92; DB 7; Length 149;
Best Local Similarity 79.2%; Pred. No. 2.2e-06;
Matches 19; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Cy 1 IELMTVIAIGILAIAPAYDY 24
    |||||:||||:||||:|||||
Db 16 IELMTVIAIGILAIAPAYDY 39

RESULT 5
US-10-467-657-3494
; Sequence 3494, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3494
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3494

Query Match          74.1%; Score 86; DB 6; Length 162;
Best Local Similarity 56.0%; Pred. No. 1.6e-05;
Matches 14; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Cy 1 IELMTVIAIGILAIAPAYDY 25
    |||||:||||:||||:|||||
Db 14 IELMTVIAIGILAIAPAYDY 38

RESULT 6
US-10-467-657-6040
; Sequence 6040, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6040
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6040

Query Match          67.2%; Score 78; DB 6; Length 129;
Best Local Similarity 56.0%; Pred. No. 0.00016;
```


Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ELMIVIAIIGILAIAPAYQDV 25
:||||:||||:||||:||||:||||:
Db 11 ELMIVIAIIGILAIAPAYSYTYI 35

RESULT 7

US-10-467-657-688
/ Sequence 688, Application US/10467657
/ Publication No. US20050260581A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MASIGNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/467,657
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ PRIOR FILING DATE: 2001-02-12
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqWin99, version 1.04
/ SEQ ID NO 688
/ LENGTH: 149
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657-688

Query Match
Best Local Similarity 55.2%; Score 64; DB 6; Length 149;
Best Local Similarity 44.0%; Pred. No. 0.016;
Matches 11; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ELMIVIAIIGILAIAPAYQDV 25
:||||:||||:||||:||||:||||:
Db 10 VELISVVLISVLAIVYPSRYNV 34

RESULT 8

US-10-467-657-3480
/ Sequence 3480, Application US/10467657
/ Publication No. US20050260581A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MASIGNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/467,657
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ PRIOR FILING DATE: 2001-02-12
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqWin99, version 1.04
/ SEQ ID NO 3480
/ LENGTH: 236
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3480

Query Match
Best Local Similarity 49.1%; Score 57; DB 6; Length 236;
Best Local Similarity 50.0%; Pred. No. 0.24;
Matches 12; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ELMIVIAIIGILAIAPAYQDV 25
:||||:||||:||||:||||:||||:
Db 28 ELMIVIAIIGILAIAPAYSGWI 51

RESULT 9

US-11-198-847-143
/ Sequence 143, Application US/11198847
/ Publication No. US20050271589A1
/ GENERAL INFORMATION:
/ APPLICANT: University of Utah Research Foundation
/ APPLICANT: Cognetix, Inc.
/ APPLICANT: Jones, Robert M.
/ APPLICANT: Garrett, James E.
/ APPLICANT: Watkins, Maren
/ APPLICANT: Oliveira, Baldoero M.
/ TITLE OF INVENTION: B-Superfamily Conotoxins
/ FILE REFERENCE: 2314-296
/ CURRENT APPLICATION NUMBER: US/11/198,847
/ CURRENT FILING DATE: 2005-08-08
/ PRIOR APPLICATION NUMBER: US 10/838,226
/ PRIOR FILING DATE: 2004-05-05
/ PRIOR APPLICATION NUMBER: US 10/058,053
/ PRIOR FILING DATE: 2000-01-29
/ PRIOR APPLICATION NUMBER: US 60/264323
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 340
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 143
/ LENGTH: 109
/ TYPE: PRT
/ ORGANISM: Conus miles
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(587)
/ OTHER INFORMATION: n may be any nucleotide
US-11-198-847-143

Query Match
Best Local Similarity 40.5%; Score 47; DB 7; Length 109;
Best Local Similarity 33.3%; Pred. No. 2.5;
Matches 7; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 3 ELMIVIAIIGILAIAPAYQDV 23
:||||:||||:||||:||||:||||:
Db 8 MMVVMMVGVTVAGSLPFPD 28

RESULT 10

US-10-467-657-6830
/ Sequence 6830, Application US/10467657
/ Publication No. US20050260581A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MASIGNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/467,657
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ PRIOR FILING DATE: 2001-02-12
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqWin99, version 1.04
/ SEQ ID NO 6830
/ LENGTH: 223
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6830

Query Match
Best Local Similarity 37.9%; Score 44; DB 6; Length 223;
Best Local Similarity 52.9%; Pred. No. 14;
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ELMIVIAIIGILAIAPAYQDV 18
:||||:||||:||||:||||:||||:
Db 149 ELAAVLLILGVAIAIAL 165

RESULT 11
US-10-878-556A-143
; Sequence 143, Application US/10878556A
; Publication No. US20050266399A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann La-Roche Inc.
; TITLE OF INVENTION: HCV regulated protein expression
; FILE REFERENCE: 21762
; CURRENT APPLICATION NUMBER: US/10/878,556A
; CURRENT FILING DATE: 2004-06-28
; NUMBER OF SEQ ID NOS: 199
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 143
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLIC INFORMATION:
; DATABASE ACCESSION NUMBER: gw_hum/call_human
; DATABASE ENTRY DATE: 1994-10-01
US-10-878-556A-143

Query Match 37.9%; Score 44; DB 6; Length 337;
Best Local Similarity 64.3%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 12 ILAIALPAYQDY 25
Db 49 ILAAQMPAYQELV 62

RESULT 12
US-10-131-826A-224
; Sequence 224, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Deenoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO: 224
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-131-826A-224

Query Match 37.9%; Score 44; DB 6; Length 449;
Best Local Similarity 32.1%; Pred. No. 30;
Matches 9; Conservative 7; Mismatches 6; Indels 6; Gaps 1;

Qy 3 LMIYVATIG-----ILAIALPAYQDY 24
Db 401 LLIVVLLGVTLFTVVLVPLQAYEST 428

RESULT 13
US-10-523-477-8
; Sequence 8, Application US/10523477
; Publication No. US20050266406A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MAXS AS MODIFIERS OF THE AXIN PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-051C-US
; CURRENT APPLICATION NUMBER: US/10/523,477
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US 60/401,534
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/411,153
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 8
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-523-477-8

Query Match 37.9%; Score 44; DB 6; Length 519;
Best Local Similarity 57.1%; Pred. No. 35;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 6 VIATIGILAIALP 19
Db 230 VLALGITLALVLP 243

RESULT 14
US-11-089-551A-2
; Sequence 2, Application US/11089551A
; Publication No. US20050266242A1
; GENERAL INFORMATION:
; APPLICANT: Lindquist et al.
; TITLE OF INVENTION: ELECTRICAL CONDUCTORS AND DEVICES FROM PRION-LIKE PROTEINS
; FILE REFERENCE: 30554/40025A
; CURRENT APPLICATION NUMBER: US/11/089,551A
; CURRENT FILING DATE: 2005-03-24
; PRIOR APPLICATION NUMBER: US 60/559,286
; PRIOR FILING DATE: 2004-03-31
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-11-089-551A-2

Query Match 37.9%; Score 44; DB 7; Length 685;
 Best Local Similarity 42.9%; Pred. No. 48;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 4 MIVAIIGIILAIAPAYDY 24
 DB 640 MIVIAVLETPAVCVETVODY 660

RESULT 15

US-10-517-939-330
 / Sequence 330, Application US/10517939
 / Publication No. US20060003433A1
 / GENERAL INFORMATION:
 / APPLICANT: Steer, Brian
 / APPLICANT: Callen, Walter
 / APPLICANT: Healey, Shaun
 / APPLICANT: Hazlewood, Geoff
 / APPLICANT: Wu, Di
 / APPLICANT: Blum, David
 / APPLICANT: Esteghlalian, Alireza
 / TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM
 / TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
 / FILE REFERENCE: 564462007901
 / CURRENT APPLICATION NUMBER: US/10/517,939
 / CURRENT FILING DATE: 2004-12-13
 / PRIOR APPLICATION NUMBER: PCT/US03/19153
 / PRIOR FILING DATE: 2003-06-16
 / PRIOR APPLICATION NUMBER: 60/389,299
 / PRIOR FILING DATE: 2002-06-14
 / NUMBER OF SEQ ID NOS: 380
 / SOFTWARE: PasteSeq for Windows Version 4.0
 / SEQ ID NO 330
 / LENGTH: 755
 / TYPE: PRT
 / ORGANISM: Unknown
 / FEATURE:
 / OTHER INFORMATION: Obtained from an environmental sample.
 / NAME/KEY: SIGNAL
 / LOCATION: (1)... (35)
 US-10-517-939-330

Query Match 37.1%; Score 43; DB 6; Length 755;
 Best Local Similarity 50.0%; Pred. No. 73;
 Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 5 IVIAIGIILAIAP 20
 DB 17 LVIALPILLAAVALPS 32

Search completed: January 23, 2006, 12:09:41
 Job time : 3.49345 secs

THE TUGO BANK (USA)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:20:59 ; Search time 82.9694 seconds
(without alignments)
132.392 Million cell updates/sec

Title: US-10-501-838a-14
Sequence: 1 IELMIVIAIGILAAIALPAYQDYV 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_21.*
2: GeneseqP1908.*
3: GeneseqP1908.*
4: GeneseqP2000.*
5: GeneseqP2001.*
6: GeneseqP2002.*
7: GeneseqP2003.*
8: GeneseqP2004.*
9: GeneseqP2005.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	25	ADBI6899	AdBI6899 Eikenella
2	116	100.0	25	ABO8231	ABO8231 Eikenella
3	115	99.1	57	AAR38501	AAR38501 M. bovis
4	112	96.6	25	ADBI6898	AdBI6898 Kingella
5	112	96.6	25	ABO8230	ABO8230 Kingella
6	111	95.7	139	AAU80946	AAU80946 Pseudomon
7	111	95.7	139	ABG94222	ABG94222 Pseudomon
8	111	95.7	139	ABG80534	ABG80534 Pseudomon
9	111	95.7	139	ADDD24109	ADDD24109 Pseudomon
10	111	95.7	139	ADU82035	ADU82035 Protein f
11	111	95.7	139	ADK17123	ADK17123 Virus-lik
12	110	94.8	50	AAP40503	AAP40503 Sequence
13	110	94.8	53	AAR38500	AAR38500 N. mening
14	110	94.8	53	AAR38499	AAR38499 N. gonorr
15	110	94.8	98	ABP71156	ABP71156 N. gonorr
16	110	94.8	111	ABE4927	ABE4927 N. gonorr
17	110	94.8	159	AAE60559	AAE60559 Proteins
18	110	94.8	167	AAV41298	AAV41298 Neisseria
19	110	94.8	167	AAU14100	AAU14100 Peptide s
20	110	94.8	167	ABO10265	ABO10265 N. gonorr
21	110	94.8	170	AAV41299	AAV41299 Neisseria
22	110	94.8	170	ABU06060	ABU06060 N. mening
23	110	94.8	170	ADP08169	ADP08169 Neisseria
24	110	94.8	170	ABE91528	ABE91528 Microbial

25	110	94.8	179	6	AAR37952
26	109	94.0	136	9	AEB39873
27	109	94.0	137	9	ABE42592
28	109	94.0	142	9	ABE36463
29	109	94.0	174	7	ABO77956
30	108	93.1	169	6	ADA34990
31	105	90.5	26	2	AAW08916
32	105	90.5	26	2	AAW61117
33	105	90.5	26	2	AAW62458
34	105	90.5	26	2	AAW75866
35	105	90.5	26	2	AAW87642
36	105	90.5	26	3	AAW71100
37	105	90.5	26	3	AAW84861
38	105	90.5	26	4	AAE10806
39	105	90.5	26	6	ADN89411
40	104	89.7	154	6	ABO10264
41	103	88.8	51	2	AAR38494
42	103	88.8	53	2	AAR38503
43	103	88.8	62	2	AAR38495
44	103	88.8	67	2	AAR38496
45	103	88.8	149	6	ABJ18747

ALIGNMENTS

RESULT 1
ADBI6899 standard; peptide; 25 AA.

ADBI6899;

20-NOV-2003 (first entry)

Bikenella corrodens Pilin penetrating peptide 14.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;
infectivity; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiant; antitartaroclerotic;
osteopathic; cytostatic; nootropic.

Eikenella corrodens.

WC2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WC-IB000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;
WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating
or preventing e.g. endocrine disorders.

Claim 2; Page 14; 60pp; English.

This invention relates to a novel peptide sequences capable of
translocating across a biological barrier. Furthermore, it refers to
methods that use these peptides to facilitate penetration of a
biologically active effector molecule such as a drug or other therapeutic
agent across biological barriers e.g. epithelial or endothelial cells
sealed by tight junctions. This peptide is derived from a bacterial
toxin, an integral membrane or extracellular protein and can comprise an
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
or enzyme. The effector molecule, however, can comprise for example
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC having neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or nocotropic activities. This peptide is from Piliin of
 CC *Elkenella* corrodens and is penetrating peptide 14 of the invention.
 XX
 SQ Sequence 25 AA;
 Query Match 100.0%; Score 116; DB 6; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.9e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 IELMIVIAIIGILAAIAPAYQDVV 25
 Db 1 IELMIVIAIIGILAAIAPAYQDVV 25
 RESULT 2
 AEB08231
 ID AEB08231 standard; peptide; 25 AA.
 AC AEB08231;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DE *Elkenella* corrodens piliin penetrating peptide 14, SEQ ID NO: 14.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW nocotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antipneumatic; cyostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS *Elkenella* corrodens.
 XX
 PN US2005136103-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 16-SEP-2004; 2004US-00942300.
 XX
 PR 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 XX
 DR WPI, 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 14; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the *Elkenella* corrodens piliin
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 SQ Sequence 25 AA;
 Query Match 100.0%; Score 116; DB 9; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.9e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 IELMIVIAIIGILAAIAPAYQDVV 25
 Db 1 IELMIVIAIIGILAAIAPAYQDVV 25
 RESULT 3
 AAR38501
 ID AAR38501 standard; protein; 57 AA.
 AC AAR38501;
 XX
 DT 25-MAR-2003 (revised)
 DT 28-OCT-1993 (first entry)
 XX
 DE M. bovis piliin protein submolecular unit.
 XX
 KW Antibodies; whole pill binding; basis; vaccine; bacterial infection;
 KW tumulant footrot infection; sheep; type IV pillated bacteria.
 XX
 OS *Moraxella* bovis.
 XX
 PN WO9311791-A1.
 XX
 PD 24-JUN-1993.
 XX
 PF 17-DEC-1992; 92WO-US011085.
 XX
 PR 18-DEC-1991; 91US-00809762.
 XX
 PA (UTOR-) UNIV OREGON HEALTH SCT.
 PA Smith AW;
 XX
 PI Smith AW;
 XX
 DR WPI, 1993-213824/26.
 XX
 PT Antigenic preparation - stimulates production of antibodies binding to
 PT piliin protein of type IV pillated bacteria, useful in vaccine compan.

PS Claim 12, Page 26, 44pp; English.

XX The sequence is that of a submolecular unit of Moraxella bovis pilin
CC protein which corresponds to at least one epitope common to structural
CC pilin proteins of Type IV pilated bacteria. It is capable of eliciting
CC antibodies (Abs) which bind to whole pill of type IV bacteria. The
CC ability of the submolecular unit to produce Abs which bind to whole pill
CC provides the basis for vaccines against type IV bacterial infections,
CC e.g. footrot infection in ruminants. (Updated on 25-MAR-2003 to correct
CC PN field.)

XX Sequence 57 AA:

Query Match 99.1%; Score 115; DB 2; Length 57;
Best Local Similarity 96.0%; Pred. No. 6.9e-10;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAITIGTIAIALPAYQDYV 25
DB 4 IELMTVIAITIGTIAIALPAYQDYI 28

RESULT 4

ADBI6898 standard; peptide; 25 AA.

ADBI6898;

20-NOV-2003 (first entry)

Kingella denitrificans pilin penetrating peptide 13.

KW Penetrating peptide; epithelial; endocervical; tight junction; diabetes;
KW infertility; hormone; vitamin deficiency; neurodegenerative;
KW cardiovascular; haematological; endocrine disorder; obesity;
KW neoplastic disease; neuroprotective; cardiac; arteriosclerotic;
KW osteopathic; cytoskeletal; neotropic.

Kingella denitrificans.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-IR000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI, 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating
or preventing e.g. endocrine disorders.

Claim 2, Page 14, 60pp; English.

XX This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and

CC endocrine disorders, as well as obesity and neoplastic disease.

CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, arteriosclerotic, osteopathic,
CC cytoskeletal or neotropic activities. This peptide is from pilin of
CC Kingella denitrificans and is penetrating peptide 13 of the invention.

XX Sequence 25 AA:

Query Match 96.6%; Score 112; DB 6; Length 25;
Best Local Similarity 96.0%; Pred. No. 7.8e-10;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAITIGTIAIALPAYQDYV 25
DB 1 IELMTVIAITIGTIAIALPAYQDYV 25

RESULT 5

ABO8230 standard; peptide; 25 AA.

ABO8230;

25-AUG-2005 (first entry)

Kingella denitrificans pilin penetrating peptide 13, SEQ ID NO: 13.

KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
KW antidiabetic; endocrine disease; gastrointestinal disease;
KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
KW neurodegenerative disease; neuroprotective; Alzheimers disease;
KW neotropic; neurological disease; parkinsons disease; antiparkinsonian;
KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
KW anticovulsant; genetic disorder; cardiovascular disease;
KW cardiovascular-gen.; atherosclerosis; arteriosclerotic;
KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
KW genitourinary disease; hematological disease; anemic;
KW autoimmune disease; immunosuppressive; immune deficiency;
KW immunostimulant; infectious disease; antimicrobial; infection;
KW erectile dysfunction; andrology; major depressive disorder;
KW antidepressant; psychiatric disorder; pain; analgesic;
KW bacterial infection; antibacterial; viral infection; virucide;
KW fungal infection; fungicide; parasitic infection; antiparasitic;
KW renal failure; antinfertility; antineumatic; cytostatic;
KW antiinflammatory; hepatotropic; hepatitis B virus infection.

Kingella denitrificans.

US2005136103-A1.

23-JUN-2005.

16-SEP-2004; 2004US-00942300.

17-SEP-2003; 2003US-00664989.

17-SEP-2003; 2003US-00665184.

17-SEP-2003; 2003US-0503615P.

(BENS/) BEN-SASSON S A.

(COHE/) COHEN E.

Ben-Sasson SA, Cohen E;

WPI, 2005-444089/45.

Composition used for translocating effectors across barrier such as
epithelial cells during treatment of e.g. endocrine disorders comprises
effector sequentially coupled with counter ion and hydrophobic agent.

Claim 46, SEQ ID NO 13, 59pp; English.

CC The present invention relates to a pharmaceutical composition of
CC penetrating peptides for transendothelial delivery of effector. The
CC invention comprises the effector sequentially coupled with a counter ion
CC and at least one hydrophobic agent, where the effector is selectively
CC encapsulated into a complex. The invention is useful for translocating
CC effectors across a biological barrier such as epithelial cells and
CC endothelial cells during treatment and prevention of disease or
CC pathological conditions (including endocrine disorders, diabetes,
CC infertility, hormone deficiencies, osteoporosis, ophthalmological
CC disorders, neurodegenerative disorders, Alzheimer's disease,
CC Parkinson's disease, multiple sclerosis, Huntington's disease,
CC cardiovascular disorder, atherosclerosis, hyper-coagulable states, hypo-
CC coagulable states, coronary disease, cerebrovascular events, metabolic
CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC hematological disorders, anemia of different entities, immunologic and
CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
CC infectious diseases, viral infections, bacterial infections, fungal
CC infections, parasitic infections, neoplastic diseases, multi-factorial
CC disorders, impotence, chronic pain, depression, different fibrosis states
CC and short stature) and for mucosal vaccination against anthrax and
CC hepatitis B. The present sequence is the Kingella dentrificans pilin
CC penetrating peptide. This sequence is used in the effective translocation
CC of amnoglycoside antibiotics and antifungal agents across an epithelial
CC barrier.

CC Sequence 25 AA;

Query Match 96.6%; Score 112; DB 9; Length 25;
Best Local Similarity 96.0%; Pred. No. 7.8e-10;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IELMIVAIIGIILAIAPAYODY 25
DB 1 IELMIVAIIGIILAIAPAYEV 25

RESULT 6
AAU80946
ID AAU80946 standard; protein; 139 AA.

AC AAU80946;

DT 09-APR-2002 (first entry)

DE Pseudomonas stutzeri pilin protein.

KM Vaccine; molecular scaffold; pilus; pilin; HBcAg; antigen;
KM hepatitis B virus capsid protein; JON; FOS; HIV gp140;
KM measles virus N protein; bee venom phospholipase; Th type 2 T-helper;
KM Th2; Sinbis virus E2 protein; amyloid beta; influenza M2 antigen;
KM human immunodeficiency virus infection; viral hepatitis; measles;
KM chicken pox; pneumonia; tuberculosis; syphilis; malaria; allergy; cancer;
KM chronic disease; arthritis; colitis; diabetes; multiple sclerosis.

OS Pseudomonas stutzeri.

XX WO200185208-A2.

PD 15-NOV-2001.

XX 02-MAY-2001; 2001WO-IB000741.

XX 05-MAY-2000; 2000US-0202341P.

PR (CYTO-) CYTOS BIOTECHNOLOGY AG.

PA (SEBB-) SEBBEL P.

PA (DUNA/) DUNANT N.

PA (BACH/) BACHMANN M.

PA (TIS/) TISSOT A.

PA (LECH/) LECHNER F.

XX Seibel P, Dunant N, Bachmann M, Tissot A, Lechener F;

DR WPI; 2002-055561/07.

XX

XX New composition, useful for vaccine production, comprises antigen or

PT antigenic determinant and non-natural molecular scaffold comprising

PT organizer and core particle such as bacterial pilus or pilin protein.

XX PS Disclosure; Page 249; 287pp; English.

CC The invention relates to a composition comprising: (a) a non-natural
CC molecular scaffold (molecular scaffold) which comprises a core particle
CC such as a bacterial pilus or pilin protein, a recombinant form of the
CC protein, a virus-like particle or a hepatitis B virus capsid protein
CC (HBcAg), and an organizer; and (b) an antigen or antigenic determinant,
CC where the molecular scaffold and antigenic determinant interact to form
CC an ordered and repetitive antigen array. Suitable antigenic determinants
CC include JON, FOS, HIV gp140, measles virus N protein, bee venom
CC phospholipase, Sinbis virus E2 protein, amyloid beta derived peptides and
CC influenza M2 antigen. The composition (or vaccine) is useful for
CC immunization, by administration to a subject, where the administration
CC produces an immune response, such as humoral, cellular or protective
CC immune response, preferably a Th type 2 T-helper (Th2) response that is
CC specific for the antigenic determinant. The administration induces
CC antibodies specific for the antigenic determinant of a subtype
CC corresponding to the Th2 subtype in the subject. The subject does not
CC generate a Th2 subtype that is specific for pilus or pilin polypeptide or
CC antigenic determinant. The composition is useful for the production of
CC vaccines for prevention of infectious diseases such as human
CC immunodeficiency virus, viral hepatitis, measles, chicken pox, pneumonia,
CC tuberculosis, syphilis, malaria, and for treating allergy, cancer, and
CC chronic diseases induced or accelerated by a Th1 type immune response,
CC such as arthritis, colitis, diabetes and multiple sclerosis. The
CC composition is useful to generate defined self-specific antibodies and
CC specific immune responses of the Th2 type and allows the creation of
CC highly efficient vaccines against infectious diseases, and for treating
CC allergy, cancer, and chronic diseases induced or accelerated by a Th1
CC type immune response. The present sequence is a peptide or protein
CC incorporated into the compositions of the invention

CC Sequence 139 AA;

Query Match 95.7%; Score 111; DB 5; Length 139;
Best Local Similarity 95.8%; Pred. No. 7.7e-09;
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IELMIVAIIGIILAIAPAYODY 24
DB 12 IELMIVAIIGIILAIAPAYDY 35

RESULT 7
ABG94222

ID ABG94222 standard; protein; 139 AA.

AC ABG94222;

DT 10-DEC-2002 (first entry)

DE Pseudomonas pilin protein.

KM Human; mouse; rat; antimicrobial; antiallergic; immunomodulatory;
KM cytostatic; antiviral; antidiabetic; hypoglycaemic; antigen array;
KM vaccine; infectious disease.

OS Pseudomonas stutzeri.

XX WO200256905-A2.

PD 25-JUL-2002.

XX 21-JAN-2002; 2002WO-IB000166.

PR 19-JAN-2001; 2001US-0262379P.

PR 04-MAY-2001; 2001US-0288549P.

PR 05-OCT-2001; 2001US-0326998P.
 PR 07-NOV-2001; 2001US-0331045P.
 XX
 XX
 XX (CYTO-) CYTOS BIOTECHNOLOGY AG.
 XX
 XX Renner WA, Bachmann M, Tlassot A, Maurer P, Lechner F, Sebbel P;
 PI Piossek C;
 XX WPI; 2002-627351/67.
 DR
 XX Molecular antigen array used in the production of vaccines for infectious
 PT diseases.
 PS
 XX Disclosure; Page 370; 441pp; English.
 CC This invention relates to a novel ordered and repetitive antigen array
 CC used in the production of vaccines for infectious diseases. The invention
 CC also discloses a composition comprising a non-natural molecular scaffold
 CC comprising a core particle selected from a core particle of a non-natural
 CC origin and a core particle of natural origin and an organiser comprising
 CC at least one first attachment site, where the organiser is connected to
 CC the core particle by at least one covalent bond. Also disclosed is an
 CC antigen or antigenic determinant with at least one second attachment
 CC site, where the antigen or antigenic determinant is amyloid beta peptide
 CC (Abeta1-42) or its fragment and where the second attachment site is
 CC selected from an attachment site not naturally occurring with the antigen
 CC or antigenic determinant and an attachment site naturally occurring with
 CC the antigen or antigenic determinant, where the second attachment site is
 CC capable of association through at least one non-peptide bond to the first
 CC attachment site and where the association to form an ordered and
 CC scaffold interact through the association to form an ordered and
 CC repetitive antigen array. The invention also comprises a coat protein
 CC capable of forming a capsid which comprises mutant Obeta coat proteins
 CC having an amino acid sequence selected from five amino acid sequences
 CC fully defined in the specification. The compounds of the invention may
 CC have antimicrobial, antiallergic, immunomodulatory, cyostatic,
 CC antiviral, antidiabetic, or hypoglycaemic activities and may be used in
 CC immunisation and as a vaccine. The present sequence represents a protein
 CC sequence used to create the compositions of the invention
 XX
 XX Sequence 139 AA;
 SQ
 Query Match 95.7%; Score 111; DB 5; Length 139;
 Best Local Similarity 95.8%; Pred. No. 7.7e-09;
 Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IELMIVAIIGIILAIAPAYQDY 24
 DB 12 IELMIVAIIGIILAIAPAYQDY 35
 RESULT 8
 ABG80534
 ID ABG80534 standard; protein; 139 AA.
 XX
 XX ABG80534;
 AC
 XX 29-NOV-2002 (first entry)
 DT
 XX Pseudomonas stutzeri pili protein.
 DE
 XX Molecular antigen array; vaccine; antigen; antimicrobial;
 KM molecular scaffold; amyloid beta; Abeta 1-42; influenza;
 KM adult respiratory distress syndrome; ARDS; Crohn's disease; anaphylaxis;
 KM graft versus host disease; Igg-mediated allergic reaction; lymphoma;
 KM allergic asthma; acute lymphoblastic leukaemia; non-Hodgkin's lymphoma;
 KM Grave's disease; systemic lupus erythematosus; osteoporosis;
 KM inflammatory immune disease; myasthenia gravis; multiple sclerosis;
 KM immunoproliferative disease; lymphadenopathy; Alzheimer's disease;
 KM angiotumproliferative lymphadenopathy; immunoblastic lymphadenopathy;
 KM rheumatoid arthritis; diabetes; infectious disease; factor Xa;
 KM enterokinase; cysteine-containing linker.
 XX

OS Pseudomonas stutzeri.
 XX
 XX WO200256907-A2.
 XX
 XX 25-JUL-2002.
 XX
 XX 21-JAN-2002; 2002MO-1B000168.
 XX
 XX 19-JAN-2001; 2001US-0262379P.
 XX
 XX 04-MAY-2001; 2001US-0286549P.
 PR 05-OCT-2001; 2001US-0326998P.
 PR 07-NOV-2001; 2001US-0331045P.
 XX
 XX (CYTO-) CYTOS BIOTECHNOLOGY AG.
 PA (NOVA) NOVAARTIS PHARMA AG.
 PA (MAUR) MAURER P.
 PA (LECH) LECHNER F.
 PA (ORTM) ORTMANN R.
 PA (LUBO) LUBEND R.
 PA (STAV) STAUFENBIEL M.
 PA (FREY) FREY P.
 PI Maurer P, Lechner F, Ortmann R, Luegend R, Staufenbiel M, Frey P;
 PI Renner WA, Bachmann M, Tlassot A, Sebbel P, Piossek C;
 DR WPI; 2002-636514/68.
 XX
 XX Molecular antigen array used in the production of vaccines for infectious
 PT diseases.
 PS
 XX Disclosure; Page 347; 418pp; English.
 CC The invention relates to a composition comprising: (a) a non-natural
 CC molecular scaffold comprising: (1) a core particle selected from: (1) a
 CC core particle of a non-natural origin; and (2) a core particle of natural
 CC origin; and (ii) an organiser comprising at least one first attachment
 CC site, where the organiser is connected to the core particle by at least
 CC one covalent bond; (b) an antigen or antigenic determinant with at least
 CC one second attachment site, where the antigen or antigenic determinant is
 CC amyloid beta peptide (Abeta 1-42) or its fragment, and where the second
 CC attachment site is selected from: (1) an attachment site not naturally
 CC occurring with the antigen or antigenic determinant; and (ii) an
 CC attachment site naturally occurring with the antigen or antigenic
 CC determinant, where the second attachment site is capable of association
 CC through at least one non-peptide bond to the first attachment site; and
 CC where the antigen or antigenic determinant and the scaffold interact
 CC through the association to form an ordered and repetitive antigen array.
 CC Also included is a process for producing a non-naturally occurring
 CC ordered and repetitive antigen array. The composition is used in
 CC immunisation and as a vaccine for diseases such as influenza, graft
 CC versus host disease, Igg-mediated allergic reactions, anaphylaxis, adult
 CC respiratory distress syndrome (ARDS), Crohn's disease, allergic asthma,
 CC acute lymphoblastic leukaemia, non-Hodgkin's lymphoma, Grave's disease,
 CC systemic lupus erythematosus, inflammatory immune diseases, myasthenia
 CC gravis, immunoproliferative disease lymphadenopathy,
 CC angiotumproliferative lymphadenopathy, immunoblastic lymphadenopathy,
 CC rheumatoid arthritis, diabetes, multiple sclerosis, Alzheimer's disease,
 CC osteoporosis and infectious diseases. The antigens are modified to possess
 CC a cleavage site (enterokinase or factor Xa) and a cysteine-containing N-
 CC or C-terminal linker peptide which serves as the attachment point to a
 CC virus like particle or bacterial protein (the scaffold protein). The
 CC present sequence is bacterial protein or peptide which is coupled to the
 CC modified antigen to form the molecular antigen array
 XX
 XX Sequence 139 AA;
 SQ
 Query Match 95.7%; Score 111; DB 5; Length 139;
 Best Local Similarity 95.8%; Pred. No. 7.7e-09;
 Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IELMIVAIIGIILAIAPAYQDY 24
 DB 12 IELMIVAIIGIILAIAPAYQDY 35

```

RESULT 9
ADD24109
ID ADD24109 standard; protein; 139 AA.
XX
XX
AC ADD24109;
XX
XX DT 15-JAN-2004 (first entry)
DE Pseudomonas stutzeri pili protein.
XX
XX vaccine composition; virus-like particle; core particle;
XX first attachment site; antigen; antigenic determinant; prion protein;
XX PrP, PrP peptide; vaccine; neuroprotective; antiinflammatory;
XX prion disease; Bovine Spongiform Encephalopathy; BSE;
XX Creutzfeldt-Jakob Disease; pili.
XX
XX Pseudomonas stutzeri.
XX
XX WO2003059386-A2.
XX
XX 24-JUL-2003.
XX
XX 17-JAN-2003; 2003WO-EP000460.
XX
XX 18-JAN-2002; 2002US-00050902.
XX 21-JAN-2002; 2002WO-1B000166.
XX 08-JUL-2002; 2002US-0393725P.
XX 18-JUL-2002; 2002US-0396590P.
XX
XX (CYTO-) CYTOS BIOTECHNOLOGY AG.
XX
XX Bachmann M, Maurer P, Pelliccioli E, Renner WA;
XX
XX WPI; 2003-598483/56.
XX
XX A vaccine composition for preventing or treating prion diseases (e.g.
XX Creutzfeldt-Jakob Disease) comprises a virus-like particle (e.g. RNA-
XX phage) and at least one prion protein or peptide bound to the virus-like
XX particle.
XX
XX Disclosure; SEQ ID NO 2; 246bp; English.
XX
XX This invention relates to a novel vaccine composition comprising a virus-
XX like or a core particle with at least one first attachment site and at
XX least one antigen or antigenic determinant that is a prion protein (PrP)
XX or its dimer, or a PrP peptide, the antigen or antigenic determinant
XX being bound to the virus-like or core particle. The vaccine of the
XX invention may have neuroprotective or antiinflammatory activity. The
XX composition is useful as a medicament or in manufacturing a medicament
XX for the treatment or prevention of prion diseases. The prion diseases may
XX include Bovine Spongiform Encephalopathy (BSE) or Creutzfeldt-Jakob
XX Disease. The present sequence is the amino acid sequence of the
XX Pseudomonas stutzeri pili protein which may be used during the creation
XX of the vaccine composition of the invention.
XX
XX Sequence 139 AA;
SQ
Query Match 95.7%; Score 111; DB 7; Length 139;
Best Local Similarity 95.8%; Pred. No. 7.7e-09;
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

XX
XX 06-MAY-2004 (first entry)
XX
XX Protein for RANKL antigen array to treat bone disease.
XX
XX osteopathic; vaccine; core particle; antigenic determinant; RANKL;
XX bone disease; encephalopathy; immune system stimulation.
XX
XX Unidentified.
XX
XX WO2003039225-A2.
XX
XX 15-MAY-2003.
XX
XX 07-NOV-2002; 2002WO-EP012449.
XX
XX 07-NOV-2001; 2001US-0331045P.
XX 18-JAN-2002; 2002US-00050902.
XX 21-JAN-2002; 2002WO-1B000166.
XX 19-JUL-2002; 2002US-0396635P.
XX
XX (CYTO-) CYTOS BIOTECHNOLOGY AG.
XX
XX Bachmann M, Maurer P, Spohn G;
XX
XX WPI; 2003-441430/41.
XX
XX New compositions comprising a core particle and at least one antigen or
XX antigenic determinant, useful for as a vaccine for therapy or prophylaxis
XX of bone diseases, particularly mammalian encephalopathies.
XX
XX Disclosure; SEQ ID NO 2; 222bp; English.
XX
XX The invention relates to a composition comprising a core particle having
XX at least one first attachment site, and at least one antigen or antigenic
XX determinant having at least one second attachment site. The antigen or
XX antigenic determinant is a RANKL protein, RANKL fragment or RANKL
XX peptide. The second attachment site is (non-) naturally occurring with
XX the antigen or antigenic determinant, and is capable of association to
XX the first attachment site. The antigen or antigenic determinant and the
XX core particle interact through the association to form an ordered and
XX repetitive antigen array. The composition is useful as a medicament, or
XX for the manufacture of a medicament for treating bone diseases. The
XX composition is especially useful for as a vaccine for therapy or
XX prophylaxis of bone diseases, particularly mammalian encephalopathies,
XX and for stimulating mammalian immune system. This sequence represents a
XX protein of the invention.
XX
XX Sequence 139 AA;
SQ
Query Match 95.7%; Score 111; DB 7; Length 139;
Best Local Similarity 95.8%; Pred. No. 7.7e-09;
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 10
ADJ82035
ID ADJ82035 standard; protein; 139 AA.
XX
XX ADJ82035;

```

```

RESULT 11
ADK17123
ID ADK17123 standard; peptide; 139 AA.
XX
XX ADK17123;
XX
XX 06-MAY-2004 (first entry)
XX
XX Virus-like particle repetitive antigen array peptide #2.
XX
XX antiallergic; antiasthmatic; cytostatic; vaccine; virus-like particle;
XX interleukin; IL-5; IL-13; eotaxin; repetitive antigen array;
XX allergic eosinophilic disease; asthma; Hodgkin's lymphoma.
XX
XX

```

OS Unidentified.
 XX MO2003040164-A2.
 PN 15-MAY-2003.
 PD 07-NOV-2002; 2002WO-EP012455.
 PP 07-NOV-2002; 2001US-0331045P.
 PR 18-JAN-2002; 2002US-00050902.
 PR 21-JAN-2002; 2002MO-1B000166.
 PR 19-JUL-2002; 2002US-035636P.
 XX (CYTO-) CYTOS BIOTECHNOLOGY AG.
 PA Bachmann M, Jemlinge G, Sonderegger I;
 PI WPI; 2003-441518/41.
 DR Composition comprising an ordered and repetitive antigen or antigenic
 XX determinant array, useful as a medicament, or for manufacturing a
 PT medicament for treating allergic eosinophilic diseases, e.g. asthma, or
 FT Hodgkin's lymphoma.
 XX Disclosure; SEQ ID NO 2; 245bp; English.
 PS The invention relates to a composition comprising a virus-like particle
 CC and at least one antigen, which is a protein or peptide of interleukin
 CC (IL)-5, IL-13 or eotaxin and is bound to the virus-like particle, or a
 CC core particle with at least one first attachment site and at least one
 CC antigen with at least one second attachment site, where the antigen is a
 CC protein or peptide of IL-5, IL-13 or eotaxin. The second attachment site
 CC is an attachment site naturally or not naturally occurring with the
 CC antigen or antigenic determinant. The second attachment site is capable
 CC of association to the first attachment site, and where the antigen or
 CC antigenic determinant and the core particle interact through the
 CC association to form an ordered and repetitive antigen array. The
 CC compositions are useful as medicaments, or for manufacturing a medicament
 CC or a vaccine for treating allergic eosinophilic diseases, e.g. asthma, or
 CC Hodgkin's lymphoma and related diseases. This sequence is used to
 CC generate the compound of the invention.
 XX
 SQ Sequence 139 AA;
 Query Match 95.7%; Score 111; DB 7; Length 139;
 Best Local Similarity 95.8%; Pred. No. 7.7e-09;
 Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IELMIVIAIIGILAAIALPAYQDY 24
 DB 12 IELMIVIAIIGILAAIALPAYQDY 35
 RESULT 12
 AAP40503
 ID AAP40503 standard; protein; 50 AA.
 XX
 AC AAP40503;
 XX
 DT 25-MAR-2003 (revised)
 DT 31-JAN-1992 (first entry)
 XX
 DE Sequence of the CNI fragment (residues 1-50) of pill from strain F62 of
 DE Neisseria gonorrhoeae.
 XX
 KM Neisseria gonorrhoeae vaccine; purified pill.
 XX
 OS Neisseria gonorrhoeae.
 XX
 FH Key Location/Qualifiers
 FT Region 1..7
 FT Modified-site 1
 FT

FT /label= N-Me-Phe
 FT Region 8..50
 FT /label= CNI
 XX
 PN US4443431-A.
 XX
 PD 17-APR-1984.
 XX
 PP 27-MAY-1981; 81US-00267538.
 XX
 PR 27-MAY-1981; 81US-00267538.
 XX
 PA (USSH) US SEC DEPT HEALTH.
 XX
 PI Buchanan TM, Pearce W, Chen KCS;
 XX
 DR WPI; 1984-113477/18.
 XX
 PT Neisseria gonorrhoeae vaccine - contg. pill fragments instead of whole
 XX pill.
 PS Example; Table 2, column 5 and 6; 6pp; English.
 XX
 CC To create the fragments of pill, pill purified according to methods
 CC described in Pearce and Buchanan (1978), or Hemmerson (1978) were
 CC reduced, carboxymethylated and then cleaved using cyanogen bromide. The
 CC largest fragment, comprising approximately 60% of the subunit of strain
 CC F62 N.9. pill, was termed CNI. The second fragment peak, approximately
 CC 32% of the subunit, was termed CNII, and a final small fragment of only 7
 CC AAs was termed CNIII. A novel Neisseria gonorrhoeae vaccine consists of a
 CC CNI fragment. (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 50 AA;
 Query Match 94.8%; Score 110; DB 1; Length 50;
 Best Local Similarity 91.7%; Pred. No. 3.4e-09;
 Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 IELMIVIAIIGILAAIALPAYQDY 24
 DB 4 IELMIVIAIIGILAAIALPAYQDY 27
 RESULT 13
 AAR38500
 ID AAR38500 standard; protein; 53 AA.
 XX
 AC AAR38500;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 28-OCT-1993 (first entry)
 XX
 DE N. meningitidis pill protein submolecular unit.
 XX
 KM Antibodies; whole pill binding; basis; vaccine; bacterial infection;
 KW ruminant footrot infection; sheep; type IV pillated bacteria.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO9311791-A1.
 XX
 PD 24-JUN-1993.
 XX
 PP 17-DEC-1992; 92WO-US011085.
 XX
 PR 18-DEC-1991; 91US-00809762.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Smith AW;
 XX
 DR WPI; 1993-213824/26.
 DR

XX Antigenic preparation - stimulates production of antibodies binding to
PT pilin protein of type IV pillated bacteria, useful in vaccine compsn.
XX
PS Claim 9; Page 25; 44pp; English.
XX
CC The sequence is that of a submolecular unit of *Neisseria meningitidis* pilin
CC protein which corresponds to at least one epitope common to structural
CC pilin proteins of Type IV pillated bacteria. It is capable of eliciting
CC antibodies (Abs) which bind to whole pill of type IV bacteria. The
CC ability of the submolecular unit to produce Abs which bind to whole pill
CC provides the basis for vaccines against type IV bacterial infections,
CC e.g. footrot infection in ruminants. (Updated on 25-MAR-2003 to correct
CC PN field.) (Updated on 27-AUG-2003 to correct OS field.)
CC
SQ Sequence 53 AA;
SQ
Query Match 94.8%; Score 110; DB 2; Length 53;
Best Local Similarity 91.7%; Pred. No. 3.7e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IELMIVIAIGIILAAALPAYQDY 24
Db 4 IELMIVIAIGIILAAALPAYQDY 27
RESULT 14
AAR38499
ID AAR38499 standard; protein; 53 AA.
XX
AC AAR38499;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 28-OCT-1993 (first entry)
XX
DE N. gonorrhea pilin protein submolecular unit.
XX
KM Antibodies, whole pill binding; basis; vaccine; bacterial infection;
KM ruminant footrot infection; sheep; type IV pillated bacteria.
XX
OS *Neisseria gonorrhoeae*.
XX
PN WO9311791-A1.
XX
PD 24-JUN-1993.
XX
PF 17-DEC-1992; 92WO-US011085.
XX
PR 18-DEC-1991; 91US-00809762.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Smith AW;
XX
DR WPI; 1993-213824/26.
XX
PT Antigenic preparation - stimulates production of antibodies binding to
PT pilin protein of type IV pillated bacteria, useful in vaccine compsn.
XX
PS Claim 6; Page 25; 44pp; English.
XX
CC The sequence is that of a submolecular unit of *Neisseria gonorrhoea* pilin
CC protein which corresponds to at least one epitope common to structural
CC pilin proteins of Type IV pillated bacteria. It is capable of eliciting
CC antibodies (Abs) which bind to whole pill of type IV bacteria. The
CC ability of the submolecular unit to produce Abs which bind to whole pill
CC provides the basis for vaccines against type IV bacterial infections,
CC e.g. footrot infection in ruminants. (Updated on 25-MAR-2003 to correct
CC PN field.) (Updated on 27-AUG-2003 to correct OS field.)
CC
SQ Sequence 53 AA;
SQ

Query Match 94.8%; Score 110; DB 2; Length 53;
Best Local Similarity 91.7%; Pred. No. 3.7e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IELMIVIAIGIILAAALPAYQDY 24
Db 4 IELMIVIAIGIILAAALPAYQDY 27
RESULT 15
ABP77156
ID ABP77156 standard; protein; 98 AA.
XX
AC ABP77156;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoea amino acid sequence SEQ ID 842.
XX
KM Antibacterial; infection; vaccine; gene therapy.
XX
OS *Neisseria gonorrhoeae*.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB002069.
XX
PR 12-FEB-2001; 2001GB-00003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Maignani V, Monaci E;
XX
DR WPI; 2003-058415/05.
DR N-PSDB; ABZ38126.
XX
PT New protein from *Neisseria gonorrhoea*, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoea infection.
XX
PS Disclosure; Page 250; 815pp; English.
XX
CC The present invention relates to proteins from *Neisseria gonorrhoea*.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoea
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 98 AA;
SQ
Query Match 94.8%; Score 110; DB 6; Length 98;
Best Local Similarity 91.7%; Pred. No. 7.4e-09;
Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 IELMIVIAIGIILAAALPAYQDY 24
Db 11 IELMIVIAIGIILAAALPAYQDY 34
Search completed: January 23, 2006, 10:28:30
Job time : 82.9694 secs


```

RESULT 2
FMW MORNO STANDARD; PRT; 154 AA.
AC P03829;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Fimbrial protein precursor (Pilin).
GN Name=tipQ;
OS Moraxella nonliquefaciens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
CX NCBI_TaxID=478;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCCTC 7784;
RX MEDLINE=92121902; PubMed=1770363;
RA Tonjum T., Marrs C.F., Rozsa F.W., Boyre K.;
RT "The type 4 pilin of Moraxella nonliquefaciens exhibits unique
RT similarities with the pilins of Neisseria gonorrhoeae and
RT Dichelobacter (Bacteroides) nodosus."
RL J. Gen. Microbiol. 137:2483-2490(1991).
RN [2]
RP PROTEIN SEQUENCE OF 7-55.
RX MEDLINE=77116125; PubMed=838045; DOI=10.1016/0014-5793(77)80008-2;
RA Prohm L.O., Sletten K.;
RT "Purification and N-terminal sequence of a fimbrial protein from
RT Moraxella nonliquefaciens."
RL FEBS Lett. 73:29-32(1977).
CC -1- SUBUNIT: The pilin are polar flexible filaments of about 5.4
CC nanometers diameter and 2.5 micrometers average length; they
CC consist of only a single polypeptide chain arranged in a helical
CC configuration of five subunits per turn in the assembled pilus.
CC -1- SIMILARITY: Belongs to the N-Me-Phe pilin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: M59707; AAA25310.1; -; Genomic_DNA.
DR PIR: A44809; A44809.
DR HSRP: P02974; 2PIL.
DR InterPro: IPR002416; Bac_GSPH.
DR InterPro: IPR012902; Methylation_N.
DR InterPro: IPR001082; Pilin.
DR InterPro: IPR001120; Prok_N_methyl_S.
DR Pfam: PF07963; N_methyl_1.
DR Pfam: PF00114; Pilin_1.
DR PRINTS: PR00885; BCTERIALGSPH.
DR PRODOM: PD000666; Pilin_1.
DR PROSITE: PS00409; PROKAR_NTER_METHYL_1.
KW Direct protein sequencing; Fimbria; Methylation.
FT PROSEP 1 6 Fimbrial protein.
FT CHAIN 7 154 N-methylphenylalanine.
FT MOD_RES 7 7
SQ SEQUENCE 154 AA; 16467 MW; 80BE4B0815563E0F CRC64;

Query Match 99.1%; Score 115; DB 1; Length 154;
Best Local Similarity 96.0%; Pred. No. 2e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IELMIVIAIIGTILAAIPAYQDYV 25
DB 10 IELMIVIAIIGTILAAIPAYQDYI 34

```

```

AC O59507;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pilin.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
CX NCBI_TaxID=476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90256243; PubMed=1971258;
RA Elleman T.C., Hoyle P.A., Lepper A.W.;
RT "Characterization of the pilin gene of Moraxella bovis Dalton 2d and
RT expression of pilin from M. bovis in Pseudomonas aeruginosa."
RL Infect. Immun. 58:1678-1684(1990).
DR EMBL: M92155; AAA53087.1; -; Genomic_DNA.
DR PIR: A41490; A41490.
DR HSRP: P02973; 10QM.
DR SMR: O59507; 7-135.
DR GO: GO:0009289; C:fimbrium; IEA.
DR GO: GO:0015627; C:type II protein secretion system complex; IEA.
DR GO: GO:0008565; F:protein transporter activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0015628; P:type II protein secretion system; IEA.
DR InterPro: IPR002416; Bac_GSPH.
DR InterPro: IPR001082; Pilin.
DR InterPro: IPR001120; Prok_N_methyl_S.
DR Pfam: PF00114; Pilin_1.
DR PRINTS: PR00885; BCTERIALGSPH.
DR PRODOM: PD000666; Pilin_1.
DR PROSITE: PS00409; PROKAR_NTER_METHYL_1.
KW Fimbria.
SQ SEQUENCE 156 AA; 15941 MW; EB281DA663B9B3C9 CRC64;

Query Match 99.1%; Score 115; DB 2; Length 156;
Best Local Similarity 96.0%; Pred. No. 2e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IELMIVIAIIGTILAAIPAYQDYV 25
DB 10 IELMIVIAIIGTILAAIPAYQDYI 34

RESULT 4
FMW MORBO STANDARD; PRT; 157 AA.
AC P07640;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Fimbrial protein Q precursor (Beta pilin) (Q pilin).
GN Name=tipQ;
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
CX NCBI_TaxID=476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EP63;
RX MEDLINE=85234350; PubMed=2861194;
RA Marrs C.F., Schoolnik G., Koomey J.M., Hardy J., Rothbard J.,
RA Falkow S.;
RT "Cloning and sequencing of a Moraxella bovis pilin gene."
RL J. Bacteriol. 163:132-139(1985).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EP63;
RX MEDLINE=90094235; PubMed=2403542;
RA Pulks K.A., Marrs C.F., Stevens S.P., Green M.R.;
RT "Sequence analysis of the inversion region containing the pilin genes
RT of Moraxella bovis."
RL J. Bacteriol. 172:310-316(1990).

```

```

RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=91286182; PubMed=2061282;
RA Roza F.W., Marrs C.F.;
RT "Interesting sequence differences between the pilin gene inversion
RL regions of Moraxella lacunata ATCC 17956 and Moraxella bovis Epp63."
RN J. Bacteriol. 173:4000-4006(1991).
[4]
RP PROTEIN SEQUENCE OF 7-157.
RX MEDLINE=89010522; PubMed=2902184; DOI=10.1084/jem.168.3.983;
RA Ruell W.W., Marrs C.F., Fernandez R., Falkow S., Schoolnik G.K.;
RT "Purification, characterization, and pathogenicity of Moraxella bovis
RN pilin."
RX J. Exp. Med. 168:983-1002(1988).
CC -1- SUBUNIT: The pilin are polar flexible filaments of about 5.4
CC nanometers diameter and 2.5 micrometers average length; they
CC consist of only a single polypeptide chain arranged in a helical
CC configuration of five subunits per turn in the assembled pilus.
CC -1- MISCELLANEOUS: Moraxella bovis can express either a Q or a I
CC pilin, the inversion of 2 kb of DNA determines which pilin is
CC expressed.
CC -1- SIMILARITY: Belongs to the N-Me-Phe pilin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; M1435; AAA5304.1; -; Genomic_DNA.
DR EMBL; M32345; AAA88223.1; -; Genomic_DNA.
DR EMBL; M59712; AAA25308.1; -; Genomic_DNA.
DR PIR; A24434; A24434.
DR HSSP; P02973; 100W.
DR InterPro; IPR000983; Bac_GSPG.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR012902; Methylation_N.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF00114; Pilin; 1.
DR PRINTS; PR00813; BCTERIALGSPG.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
KW Direct protein sequencing; Fimbria; Methylation.
FT PROSP 1
FT CHAIN 1 157 Fimbrial protein Q.
FT MOD_RES 7 7 N-methylphenylalanine.
FT DISTU_FID 136 155 By similarity.
SQ SEQUENCE 157 AA; 16006 MW; A923CD8A26C693C9 CRC64;

Query Match 99.1%; Score 115; DB 1; Length 157;
Best Local Similarity 96.0%; Pred. No. 2e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIITGILAAALPAYODVY 25
DB 10 IELMIVAIITGILAAALPAYODYI 34

RESULT 5
059503 MORBO PRELIMINARY; PRT; 157 AA.
AC Q59503;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Prepilin.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
NCBI_TaxID=476;

```

```

RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=FL462 serogroup G;
RC MEDLINE=94327452; PubMed=8051000;
RA Atwell J.L., Tennent J.M., Lepper A.W., Eilerman T.C.;
RT "Characterization of pilin genes from seven serologically defined
RL prototype strains of Moraxella bovis."
RN J. Bacteriol. 176:4875-4882(1994).
DR EMBL; L32966; AAA53559.1; -; Genomic_DNA.
DR HSSP; P02973; 100W.
DR GO; GO:0009289; C:fimbrium; IEA.
DR GO; GO:0009289; C:protein secretion system complex; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; P:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR000983; Bac_GSPG.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF00114; Pilin; 1.
DR PRINTS; PR00813; BCTERIALGSPG.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria.
SQ SEQUENCE 157 AA; 16327 MW; BE4FCA0F382430D2 CRC64;

Query Match 99.1%; Score 115; DB 2; Length 157;
Best Local Similarity 96.0%; Pred. No. 2e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIITGILAAALPAYODVY 25
DB 10 IELMIVAIITGILAAALPAYODYI 34

RESULT 6
059501 MORBO PRELIMINARY; PRT; 158 AA.
AC Q59501;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Prepilin.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
NCBI_TaxID=476;
[1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=218R serogroup F;
RC MEDLINE=94327452; PubMed=8051000;
RA Atwell J.L., Tennent J.M., Lepper A.W., Eilerman T.C.;
RT "Characterization of pilin genes from seven serologically defined
RL prototype strains of Moraxella bovis."
RN J. Bacteriol. 176:4875-4882(1994).
DR EMBL; L32965; AAA53558.1; -; Genomic_DNA.
DR PIR; A55851; A55851.
DR HSSP; P02973; 100W.
DR GO; GO:0009289; C:fimbrium; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; P:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR000983; Bac_GSPG.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF00114; Pilin; 1.
DR PRINTS; PR00813; BCTERIALGSPG.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.

```

```

KW Fimbria.
SQ SEQUENCE 158 AA; 16216 MW; 620B513CA4DC39A5 CRC64;

Query Match
Best Local Similarity 99.1%; Score 115; DB 2; Length 158;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IELMTVIAIIGILAIAPAYDYY 25
Db 10 IELMTVIAIIGILAIAPAYDYY 34

RESULT 7
059508 MORBO PRELIMINARY; PRT; 158 AA.
AC Q59508;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Prepilin.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxID=476;

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3W07 serogroup B;
RX MEDLINE=94327452; PubMed=8051000;
RA Atwell J.L., Tennent J.M., Lepper A.W., Elleman T.C.;
RT "Characterization of pilin genes from seven serologically defined
RT prototype strains of Moraxella bovis.";
RL J. Bacteriol. 176:4875-4882(1994).
DR EMBL, J32969; AA53565.1; -, Genomic DNA.
DR HSSP, P02973; 100W.
DR GO; GO:0009289; C:fimbrium; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR000983; Bac_GSPG.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF00114; Pilin; 1.
DR PRINTS; PR00813; BCTERIALGSPG.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria.
SQ SEQUENCE 158 AA; 16590 MW; 5958DC65678051F6 CRC64;

Query Match
Best Local Similarity 99.1%; Score 115; DB 2; Length 158;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IELMTVIAIIGILAIAPAYDYY 25
Db 10 IELMTVIAIIGILAIAPAYDYY 34

RESULT 8
FMI MORBO STANDARD; PRT; 159 AA.
AC P20657;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Fimbrial protein I precursor (Alpha pilin) (I pilin).
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxID=476;
RX [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RRP63;
RX MEDLINE=90094235; PubMed=2403542;
RA Fulkis K.A., Marrs C.F., Stevens S.P., Green M.R.;
RT "Sequence analysis of the inversion region containing the pilin genes
RT of Moraxella bovis.";
RL J. Bacteriol. 172:310-316(1990).
RN [2]
RP PROTEIN SEQUENCE OF 7-159.
RX MEDLINE=89010522; PubMed=2902184; DOI=10.1084/jem.168.3.983;
RA Ruehl W.W., Marrs C.F., Fernandez R., Falkow S., Schoolnik G.K.;
RT "Purification, characterization, and pathogenicity of Moraxella bovis
RT pilin.".
RL J. Exp. Med. 168:983-1002(1988).
CC -1- SUBUNIT: The pilin are polar flexible filaments of about 5.4
CC nanometers diameter and 2.5 micrometers average length; they
CC consist of only a single polypeptide chain arranged in a helical
CC configuration of five subunits per turn in the assembled pilus.
CC -1- MISCELLANEOUS: Moraxella bovis can express either a Q or a I
CC pilin, the inversion of 2 kb of DNA determines which pilin is
CC expressed.
CC -1- SIMILARITY: Belongs to the N-Me-Phe pilin family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, M32345; -, NOT_ANNOTATED_CDS; Genomic DNA.
CC PIR, J10071; J10071.
CC HSSP, P02973; 100W.
CC InterPro; IPR002416; Bac_GSPH.
CC InterPro; IPR012902; Methylation_N.
CC InterPro; IPR001082; Pilin.
CC InterPro; IPR001120; Prok_N_methyl_S.
CC Pfam; PF07963; N_methyl; 1.
CC Pfam; PF00114; Pilin; 1.
CC PRINTS; PR00885; BCTERIALGSPH.
CC PRODOM; PD000666; Pilin; 1.
CC PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
CC Direct protein sequencing; Fimbria; Methylation.
CC PROPEP 1
CC CHAIN 7 159 Fimbrial protein I.
CC MOD_RES 7 7 N-methylphenylalanine.
CC CONFLICT 159 159 K -> KSK (in Ref. 2).
SQ SEQUENCE 159 AA; 16723 MW; 9130E2289C7F6798 CRC64;

Query Match
Best Local Similarity 99.1%; Score 115; DB 1; Length 159;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IELMTVIAIIGILAIAPAYDYY 25
Db 10 IELMTVIAIIGILAIAPAYDYY 34

RESULT 9
Q60163 MORBO PRELIMINARY; PRT; 159 AA.
AC Q60163;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Prepilin.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxID=476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RS93L serotype D, and H356CS serogroup D;
RX MEDLINE=94327452; PubMed=8051000;

```



```

RA Atwell J.L., Tennent J.M., Lepper A.W., Elleman T.C.;
RT "Characterization of pilin genes from seven serologically defined
RL J. Bacteriol. 176:4875-4882(1994).
DR EMBL, U32971; AAA53563.1; -; Genomic_DNA.
DR EMBL, U32970; AAA53562.1; -; Genomic_DNA.
DR HSSP, P02973; 100W.
DR GO, GO:0009289; C:filbrinium; IEA.
DR GO, GO:0015627; C:type II protein secretion system complex; IEA.
DR GO, GO:0008565; F:protein transporter activity; IEA.
DR GO, GO:0007155; P:cell adhesion; IEA.
DR GO, GO:0015628; P:type II protein secretion system; IEA.
DR InterPro, IPR000983; Bac_GSPG.
DR InterPro, IPR002416; Bac_GSPH.
DR InterPro, IPR001082; Pilin.
DR InterPro, IPR001120; Prok_N_methyl_S.
DR Pfam, PF00114; Pilin; 1.
DR PRINTS, PR00813; BACTERIALGSPG.
DR PRODOM, PD000666; Pilin; 1.
DR PROSITE, PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria.
SQ
SEQUENCE 159 AA; 16459 MW; P2113DD7ECD578FC CRC64;

Query Match 99.1%; Score 115; DB 2; Length 159;
Best Local Similarity 96.0%; Pred. No. 2e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITIGILAAIALPAYQDYV 25
Db 10 IELMIVIAITIGILAAIALPAYQDYI 34

RESULT 10
ID Q59505 MORBO PRELIMINARY; PRT; 159 AA.
AC Q59505;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prepilin.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxID=476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TAT849 serogroup E;
RX MEDLINE=94327452; PubMed=8051000;
RA Atwell J.L., Tennent J.M., Lepper A.W., Elleman T.C.;
RT "Characterization of pilin genes from seven serologically defined
RL J. Bacteriol. 176:4875-4882(1994).
DR EMBL, U32972; AAA53561.1; -; Genomic_DNA.
DR HSSP, P02973; 100W.
DR GO, GO:0009289; C:filbrinium; IEA.
DR GO, GO:0015627; C:type II protein secretion system complex; IEA.
DR GO, GO:0008565; F:protein transporter activity; IEA.
DR GO, GO:0007155; P:cell adhesion; IEA.
DR GO, GO:0015628; P:type II protein secretion system; IEA.
DR InterPro, IPR000983; Bac_GSPG.
DR InterPro, IPR002416; Bac_GSPH.
DR InterPro, IPR001082; Pilin.
DR InterPro, IPR001120; Prok_N_methyl_S.
DR Pfam, PF00114; Pilin; 1.
DR PRINTS, PR00813; BACTERIALGSPG.
DR PRODOM, PD000666; Pilin; 1.
DR PROSITE, PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria.
SQ
SEQUENCE 159 AA; 16459 MW; P2113DD7ECD578FC CRC64;

```

```

Best Local Similarity 96.0%; Pred. No. 2e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITIGILAAIALPAYQDYV 25
Db 10 IELMIVIAITIGILAAIALPAYQDYI 34

RESULT 11
ID Q59504 MORBO PRELIMINARY; PRT; 160 AA.
AC Q59504;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prepilin.
OS Moraxella bovis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxID=476;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=S276R serogroup A;
RX MEDLINE=94327452; PubMed=8051000;
RA Atwell J.L., Tennent J.M., Lepper A.W., Elleman T.C.;
RT "Characterization of pilin genes from seven serologically defined
RL J. Bacteriol. 176:4875-4882(1994).
DR EMBL, U32968; AAA53560.1; -; Genomic_DNA.
DR HSSP, P02973; 100W.
DR GO, GO:0009289; C:filbrinium; IEA.
DR GO, GO:0015627; C:type II protein secretion system complex; IEA.
DR GO, GO:0008565; F:protein transporter activity; IEA.
DR GO, GO:0007155; P:cell adhesion; IEA.
DR GO, GO:0015628; P:type II protein secretion system; IEA.
DR InterPro, IPR000983; Bac_GSPG.
DR InterPro, IPR002416; Bac_GSPH.
DR InterPro, IPR001082; Pilin.
DR InterPro, IPR001120; Prok_N_methyl_S.
DR Pfam, PF00114; Pilin; 1.
DR PRINTS, PR00813; BACTERIALGSPG.
DR PRODOM, PD000666; Pilin; 1.
DR PROSITE, PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria.
SQ
SEQUENCE 160 AA; 16548 MW; B39939B7A10BF20B CRC64;

Query Match 99.1%; Score 115; DB 2; Length 160;
Best Local Similarity 96.0%; Pred. No. 2e-07;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVIAITIGILAAIALPAYQDYV 25
Db 10 IELMIVIAITIGILAAIALPAYQDYI 34

RESULT 12
ID Q9RAK1 BIKCO PRELIMINARY; PRT; 32 AA.
AC Q9RAK1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Type IV pilin (fragment).
OS Bifrenella corrodens.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Bifrenella.
OX NCBI_TaxID=539;
RN [1]
RP PROTEIN SEQUENCE.
RC MEDLINE=95369934; PubMed=7642307;
RA Hood B.L., Hirschberg R.;
RT Infect. Immun. 63:3693-3696(1995).

```

```

DR GO; GO:0009289; C:fimbrium; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
SQ SEQUENCE 32 AA; 3477 MW; D54B8E058D76E5AB CRC64;

Query Match          96.6%; Score 112; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTIAIALPAYODY 24
DB 4 IELMTVIAIIGTIAIALPAYODY 27

RESULT 13
P78000_9NEIS PRELIMINARY; PRT; 144 AA.
AC P78000;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE KdpB (kdpA) (kdpC).
GN Name=kdpB; Synonyms=kdpA, kdpC;
OS Kingella denitrificans.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Kingella.
CX NCBI_TaxID=502;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33394;
RX MEDLINE=97101010; PubMed=8945537;
RA Weir S., Lee L.W., Mairs C.F.;
RT "Identification of four complete type 4 pilin genes in a single
   Kingella denitrificans genome.";
   Infect. Immun. 64:4993-4999(1996).
DR EMBL; U68761; AAB48593.1; -; Genomic DNA.
DR EMBL; U68760; AAB48592.1; -; Genomic DNA.
DR HSSP; P02974; 1AY2
DR GO; GO:0009289; C:fimbrium; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF00114; Pilin; 1.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria.
SQ SEQUENCE 144 AA; 14718 MW; 849190D02P92A0AF CRC64;

Query Match          96.6%; Score 112; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.7e-07;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTIAIALPAYODY 24
DB 11 IELMTVIAIIGTIAIALPAYODY 34

RESULT 14
O85826_EIKCO PRELIMINARY; PRT; 153 AA.
ID O85826_EIKCO PRELIMINARY; PRT; 153 AA.
AC O85826;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Type IV pilin.
GN Name=pilA1;
OS Bifrenella corrodens.
OS Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

```

```

OC Neisseriaceae; Bifrenella.
CX NCBI_TaxID=539;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VAL;
RX MEDLINE=99328958; PubMed=10400570;
RA Villar M.T., Helber J.T., Hood B., Schaefer M.R., Hirschberg R.L.;
RT "Bifrenella corrodens phase variation involves a posttranslational
   event in pilus formation.";
   J. Bacteriol. 181:4154-4160(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VAL;
RA Villar M., Helber J.T., Hood B., Schaefer M.R., Hirschberg R.L.;
RX Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF079304; AAC28468.1; -; Genomic DNA.
DR HSSP; P02974; 2PTL.
DR GO; GO:0009289; C:fimbrium; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR001082; Pilin.
DR Pfam; PF00114; Pilin; 1.
DR PRINTS; PR00885; BCTERIALGSPH.
DR PRODOM; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria.
SQ SEQUENCE 153 AA; 15662 MW; BACC962C59F85A9B CRC64;

Query Match          96.6%; Score 112; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 5e-07;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAIIGTIAIALPAYODY 24
DB 12 IELMTVIAIIGTIAIALPAYODY 35

RESULT 15
P77876_9NEIS PRELIMINARY; PRT; 156 AA.
ID P77876_9NEIS PRELIMINARY; PRT; 156 AA.
AC P77876;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE KdpD.
GN Name=kdpD;
OS Kingella denitrificans.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Kingella.
CX NCBI_TaxID=502;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33394;
RX MEDLINE=97101010; PubMed=8945537;
RA Weir S., Lee L.W., Mairs C.F.;
RT "Identification of four complete type 4 pilin genes in a single
   Kingella denitrificans genome.";
   Infect. Immun. 64:4993-4999(1996).
DR EMBL; U68761; AAB48594.1; -; Genomic DNA.
DR HSSP; P02974; 2PTL.
DR GO; GO:0009289; C:fimbrium; IEA.
DR GO; GO:0015627; C:type II protein secretion system complex; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0015628; P:type II protein secretion system; IEA.
DR InterPro; IPR002416; Bac_GSPH.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR001120; Prok_N_methyl_S.
DR Pfam; PF00114; Pilin; 1.

```

DR PRINTS; PR00885; BCTERIALGSPH.
 DR Prodom; PD000666; P11n; 1.
 DR PROSITS; PS00409; PROKAR_NTER_METHYL; 1.
 KW P11n; 1.
 SQ SEQUENCE 156 AA; 16766 MW; BE050AFED0B56A99 CRC64;

Query Match 96.6%; Score 112; DB 2; Length 156;
 Best Local Similarity 96.0%; Pred. No. 5e-07;
 Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 HELMIVIAIGIILAIAPAYQDYV 25
 |||||
 DB 15 HELMIVIAIGIILAIAPAYQDYV 39

Search completed: January 23, 2006, 10:40:53
 Job time : 79.6943 secs

Blank (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:52:43 ; Search time 57.4236 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-14
Perfect score: 116
Sequence: 1 IELMIVAIIGILAAIALPAYQDVV 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications_A1_Main:
1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubppa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	25	4	US-10-665-184-14
2	116	100.0	25	5	US-10-942-300-14
3	116	100.0	25	5	US-10-501-838a-14
4	112	96.6	25	4	US-10-665-184-13
5	112	96.6	25	5	US-10-942-300-13
6	112	96.6	25	5	US-10-501-838a-13
7	111	95.7	139	3	US-09-848-616-140
8	111	95.7	139	4	US-10-289-454-2
9	111	95.7	139	4	US-10-050-902-140
10	111	95.7	139	4	US-10-050-898-140
11	111	95.7	139	4	US-10-346-190-2
12	111	95.7	139	4	US-10-289-456-2
13	110	94.8	167	4	US-10-267-682-109
14	110	94.8	167	4	US-10-267-748-109
15	110	94.8	179	5	US-10-472-260-130
16	110	94.8	179	5	US-10-497-846-4
17	105	90.5	26	3	US-09-086-118-29
18	105	90.5	26	3	US-09-086-118-29
19	105	90.5	26	3	US-09-086-118-29
20	105	90.5	26	3	US-09-086-118-29
21	105	90.5	26	3	US-09-086-118-29
22	105	90.5	26	3	US-09-086-118-29
23	105	90.5	26	3	US-09-086-118-29
24	104	89.7	154	4	US-10-267-682-108
25	104	89.7	154	4	US-10-267-748-108
26	103	88.8	149	4	US-10-127-032-96
27	92	79.3	149	6	US-11-019-005-2

28	92	79.3	149	6	US-11-019-005-26	Sequence 26, Appl
29	92	79.3	149	6	US-11-019-005-28	Sequence 28, Appl
30	92	79.3	149	6	US-11-019-005-30	Sequence 30, Appl
31	92	79.3	149	6	US-11-019-005-32	Sequence 32, Appl
32	92	79.3	149	6	US-11-019-005-34	Sequence 34, Appl
33	92	79.3	149	6	US-11-019-005-38	Sequence 38, Appl
34	92	79.3	149	6	US-11-019-005-40	Sequence 40, Appl
35	92	79.3	149	6	US-11-019-005-42	Sequence 42, Appl
36	92	79.3	149	6	US-11-019-005-44	Sequence 44, Appl
37	88	75.9	145	6	US-11-019-005-36	Sequence 36, Appl
38	87	75.0	149	6	US-10-282-122a-56078	Sequence 56078, A
39	84	72.4	292	3	US-09-975-719-419	Sequence 419, App
40	59	50.9	213	3	US-09-975-719-383	Sequence 383, App
41	57	49.1	288	5	US-10-450-763-55321	Sequence 55321, A
42	54	46.6	214	4	US-10-389-566-616	Sequence 616, App
43	53.5	46.1	289	4	US-10-437-963-195929	Sequence 195929, Sequence
44	53	45.7	159	4	US-10-437-963-188523	Sequence 188523, Sequence
45	51.5	44.4	274	5	US-10-994-726-190	Sequence 190, App

ALIGNMENTS

RESULT 1
US-10-665-184-14
Sequence 14, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 14
LENGTH: 25
TYPE: PRT
ORGANISM: Elkenella corrodens
US-10-665-184-14

Query Match 100.0%; Score 116; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMIVAIIGILAAIALPAYQDVV 25
DB 1 IELMIVAIIGILAAIALPAYQDVV 25

RESULT 2
US-10-942-300-14
Sequence 14, Application US/10942300
Publication No. US2005016103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-503
CURRENT APPLICATION NUMBER: US/10/942,300
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 14
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Eikenella corrodens
US-10-942-300-14

Query Match 100.0%; Score 116; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 1,3e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TELMIVIAIIGTIAIALPAYQDYV 25
Db 1 TELMIVIAIIGTIAIALPAYQDYV 25

RESULT 3
US-10-501-838A-14
; Sequence 14, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Eikenella corrodens
US-10-501-838A-14

Query Match 100.0%; Score 116; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 1,3e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TELMIVIAIIGTIAIALPAYQDYV 25
Db 1 TELMIVIAIIGTIAIALPAYQDYV 25

RESULT 4
US-10-665-184-13
; Sequence 13, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Kingella denitrificans

US-10-665-184-13

Query Match 96.6%; Score 112; DB 4; Length 25;
Best Local Similarity 96.0%; Pred. No. 4,8e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TELMIVIAIIGTIAIALPAYQDYV 25
Db 1 TELMIVIAIIGTIAIALPAYQDYV 25

RESULT 5
US-10-942-300-13
; Sequence 13, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT FILING DATE: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Kingella denitrificans
US-10-942-300-13

Query Match 96.6%; Score 112; DB 5; Length 25;
Best Local Similarity 96.0%; Pred. No. 4,8e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TELMIVIAIIGTIAIALPAYQDYV 25
Db 1 TELMIVIAIIGTIAIALPAYQDYV 25

RESULT 6
US-10-501-838A-13
; Sequence 13, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 13
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Kingella denitrificans
US-10-501-838A-13

Query Match 96.6%; Score 112; DB 5; Length 25;
Best Local Similarity 96.0%; Pred. No. 4,8e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy	1	I E L M V I A I G I L T A A I A L P A Y Q D Y	25
		:	
Db	1	I E L M V I A I G I L T A A I A L P A Y Q E Y V	25

```

RESULT 7
US-09-848-616-140
: Sequence 140, Application US/09848616
: Publication No. US20030054010A1
: GENERAL INFORMATION:
: APPLICANT: Seibel, Peter
: APPLICANT: Dunant, Nicolas
: APPLICANT: Bachmann, Martin
: APPLICANT: Tisot, Alain
: APPLICANT: Lechner, Franziska
: TITLE OF INVENTION: Molecular Antigen Array
: FILE REFERENCE: 1700.018002
: CURRENT APPLICATION NUMBER: US/09/848.616
: CURRENT FILING DATE: 2001-05-05
: NUMBER OF SEQ ID NOS: 186
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 140
: LENGTH: 139
: TYPE: PRT
: ORGANISM: Pseudomonas stutzeri
US-09-848-616-140

```

```
Query Match      95.7%;      Score 111;      DB 3;      Length 139
Best Local Similarity 95.8%;      Pred. NO. 4.4e-08;
Matches      23;      Conservative      1;      Mismatches      0;      Indels
```

```

Oy      1 IELMIVAIIGILAAIALPAYQDY 24
         |||||:|||||
Db      12 IELMIVAIIGILAAIALPAYQDY 35

```

```

      8
      9
      10
      11
      12
      13
      14
      15
      16
      17
      18
      19
      20
      21
      22
      23
      24
      25
      26
      27
      28
      29
      30
      31
      32
      33
      34
      35
      36
      37
      38
      39
      40
      41
      42
      43
      44
      45
      46
      47
      48
      49
      50
      51
      52
      53
      54
      55
      56
      57
      58
      59
      60
      61
      62
      63
      64
      65
      66
      67
      68
      69
      70
      71
      72
      73
      74
      75
      76
      77
      78
      79
      80
      81
      82
      83
      84
      85
      86
      87
      88
      89
      90
      91
      92
      93
      94
      95
      96
      97
      98
      99
      100
      101
      102
      103
      104
      105
      106
      107
      108
      109
      110
      111
      112
      113
      114
      115
      116
      117
      118
      119
      120
      121
      122
      123
      124
      125
      126
      127
      128
      129
      130
      131
      132
      133
      134
      135
      136
      137
      138
      139
      140
      141
      142
      143
      144
      145
      146
      147
      148
      149
      150
      151
      152
      153
      154
      155
      156
      157
      158
      159
      160
      161
      162
      163
      164
      165
      166
      167
      168
      169
      170
      171
      172
      173
      174
      175
      176
      177
      178
      179
      180
      181
      182
      183
      184
      185
      186
      187
      188
      189
      190
      191
      192
      193
      194
      195
      196
      197
      198
      199
      200
      201
      202
      203
      204
      205
      206
      207
      208
      209
      210
      211
      212
      213
      214
      215
      216
      217
      218
      219
      220
      221
      222
      223
      224
      225
      226
      227
      228
      229
      230
      231
      232
      233
      234
      235
      236
      237
      238
      239
      240
      241
      242
      243
      244
      245
      246
      247
      248
      249
      250
      251
      252
      253
      254
      255
      256
      257
      258
      259
      260
      261
      262
      263
      264
      265
      266
      267
      268
      269
      270
      271
      272
      273
      274
      275
      276
      277
      278
      279
      280
      281
      282
      283
      284
      285
      286
      287
      288
      289
      290
      291
      292
      293
      294
      295
      296
      297
      298
      299
      300
      301
      302
      303
      304
      305
      306
      307
      308
      309
      310
      311
      312
      313
      314
      315
      316
      317
      318
      319
      320
      321
      322
      323
      324
      325
      326
      327
      328
      329
      330
      331
      332
      333
      334
      335
      336
      337
      338
      339
      340
      341
      342
      343
      344
      345
      346
      347
      348
      349
      350
      351
      352
      353
      354
      355
      356
      357
      358
      359
      360
      361
      362
      363
      364
      365
      366
      367
      368
      369
      370
      371
      372
      373
      374
      375
      376
      377
      378
      379
      380
      381
      382
      383
      384
      385
      386
      387
      388
      389
      390
      391
      392
      393
      394
      395
      396
      397
      398
      399
      400
      401
      402
      403
      404
      405
      406
      407
      408
      409
      410
      411
      412
      413
      414
      415
      416
      417
      418
      419
      420
      421
      422
      423
      424
      425
      426
      427
      428
      429
      430
      431
      432
      433
      434
      435
      436
      437
      438
      439
      440
      441
      442
      443
      444
      445
      446
      447
      448
      449
      450
      451
      452
      453
      454
      455
      456
      457
      458
      459
      460
      461
      462
      463
      464
      465
      466
      467
      468
      469
      470
      471
      472
      473
      474
      475
      476
      477
      478
      479
      480
      481
      482
      483
      484
      485
      486
      487
      488
      489
      490
      491
      492
      493
      494
      495
      496
      497
      498
      499
      500
      501
      502
      503
      504
      505
      506
      507
      508
      509
      510
      511
      512
      513
      514
      515
      516
      517
      518
      519
      520
      521
      522
      523
      524
      525
      526
      527
      528
      529
      530
      531
      532
      533
      534
      535
      536
      537
      538
      539
      540
      541
      542
      543
      544
      545
      546
      547
      548
      549
      550
      551
      552
      553
      554
      555
      556
      557
      558
      559
      560
      561
      562
      563
      564
      565
      566
      567
      568
      569
      570
      571
      572
      573
      574
      575
      576
      577
      578
      579
      580
      581
      582
      583
      584
      585
      586
      587
      588
      589
      590
      591
      592
      593
      594
      595
      596
      597
      598
      599
      600
      601
      602
      603
      604
      605
      606
      607
      608
      609
      610
      611
      612
      613
      614
      615
      616
      617
      618
      619
      620
      621
      622
      623
      624
      625
      626
      627
      628
      629
      630
      631
      632
      633
      634
      635
      636
      637
      638
      639
      640
      641
      642
      643
      644
      645
      646
      647
      648
      649
      650
      651
      652
      653
      654
      655
      656
      657
      658
      659
      660
      661
      662
      663
      664
      665
      666
      667
      668
      669
      670
      671
      672
      673
      674
      675
      676
      677
      678
      679
      680
      681
      682
      683
      684
      685
      686
      687
      688
      689
      690
      691
      692
      693
      694
      695
      696
      697
      698
      699
      700
      701
      702
      703
      704
      705
```

Query Match	95.74;	Score 111;	DB 4;	Length 139;
Best Local Similarity	95.84;	Pred. No. 4.4e-08;		
Matches	23;	Conservative	1;	Mismatches 0;
			Indels	0;
			Gaps	0

```

Oy      1 1EIMVIAIIGILAAIALPAYQDY 24
         |||||:|||||
Db      12 1EIMVIAIIGILAAIALPAYQDY 35

```

RESULT 9
US-10-050-902-140
; Sequence 140, Application US/10050902
; Publication No. US20030175290A1
; GENERAL INFORMATION:
; Address: Dallas, TX 75201

Query Match	95.7%	Score 111;	DB 4;	Length 139;
Best Local Similarity	95.8%	Pred. NO. 4.4e-08;		
Matches 23; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

```

Qy      1 IELMIVAIIGILAAIALPAYQDY 24
         |||||:|||||
Db      12 IELMIVAIIGILAAIALPAYQDY 35

```

```

RESULT 10
US-10-050-898-140
/ Sequence 140, Application US/10050898
/ Publication No. US2003017571A1
/ GENERAL INFORMATION:
APPLICANT: Renner, Wolfgang A.
APPLICANT: Bachmann, Martin
APPLICANT: Tissot, Alain
APPLICANT: Meurer, Patrick
APPLICANT: Lechner, Franziska
APPLICANT: Seibel, Peter
APPLICANT: Ploessek, Christine
APPLICANT: Ottmann, Rainer
APPLICANT: Luond, Rainer
APPLICANT: Straufenbiel, Matthias
APPLICANT: Frey, Peter
/ TITLE OF INVENTION: Molecular Antigen Array
/ FILE REFERENCE: 1700, 0190005
/ CURRENT APPLICATION NUMBER: US/10/050, 898
/ PRIOR FILING DATE: 2002-01-18
/ PRIOR APPLICATION NUMBER: US 60/262, 379
/ PRIOR FILING DATE: 2001-01-19
/ PRIOR APPLICATION NUMBER: US 60/288, 549
/ PRIOR FILING DATE: 2001-05-04
/ PRIOR APPLICATION NUMBER: US 60/326, 998
/ PRIOR FILING DATE: 2001-10-05
/ PRIOR APPLICATION NUMBER: US 60/331, 045
/ PRIOR FILING DATE: 2001-11-07
/ NUMBER OF SEQ ID NOS: 350
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 140

```

LENGTH: 139
TYPE: PRT
ORGANISM: Pseudomonas stutzeri
US-10-050-898-140

Query Match 95.7%; Score 111; DB 4; Length 139;
Best Local Similarity 95.8%; Pred. No. 4.4e-08;
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAIAPAYDY 24
Db 12 IELMIVAIIGILAIAPAYDY 35

RESULT 11
US-10-346-190-2
Sequence 2, Application US/10346190
Publication No. US20030219459A1
GENERAL INFORMATION:
APPLICANT: Bachmann, Martin
APPLICANT: Maurer, Patrick
APPLICANT: Pellisoli, Erica
APPLICANT: Renner, Wolfgang A.
TITLE OF INVENTION: Prion Protein Carrier-Conjugates
FILE REFERENCE: 1700.0290003
CURRENT APPLICATION NUMBER: US/10/346,190
CURRENT FILING DATE: 2003-01-17
PRIOR APPLICATION NUMBER: 60/396,590
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/393,725
PRIOR FILING DATE: 2002-07-08
PRIOR APPLICATION NUMBER: 60/389,898
PRIOR FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: PCT/IB02/00166
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: 10/050,902
PRIOR FILING DATE: 2002-01-18
NUMBER OF SEQ ID NOS: 164
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 139
TYPE: PRT
ORGANISM: Pseudomonas stutzeri
US-10-346-190-2

Query Match 95.7%; Score 111; DB 4; Length 139;
Best Local Similarity 95.8%; Pred. No. 4.4e-08;
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAIAPAYDY 24
Db 12 IELMIVAIIGILAIAPAYDY 35

RESULT 12
US-10-289-456-2
Sequence 2, Application US/10289456
Publication No. US20040033211A1
GENERAL INFORMATION:
APPLICANT: Bachmann, Martin
APPLICANT: Maurer, Patrick
APPLICANT: Spohn, Gunther
TITLE OF INVENTION: Antigen Arrays for Treatment of Bone Disease
FILE REFERENCE: 1700.0330001
CURRENT APPLICATION NUMBER: US/10/289,456
CURRENT FILING DATE: 2002-11-07
PRIOR APPLICATION NUMBER: PCT/IB02/00166
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: US 10/050,902
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/396,635
PRIOR FILING DATE: 2002-07-19
PRIOR APPLICATION NUMBER: US 60/331,045

PRIOR FILING DATE: 2001-11-07
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 139
TYPE: PRT
ORGANISM: Pseudomonas stutzeri
US-10-289-456-2

Query Match 95.7%; Score 111; DB 4; Length 139;
Best Local Similarity 95.8%; Pred. No. 4.4e-08;
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IELMIVAIIGILAIAPAYDY 24
Db 12 IELMIVAIIGILAIAPAYDY 35

RESULT 13
US-10-267-682-109
Sequence 109, Application US/10267682
Publication No. US20040033235A1
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
Matchews, Thomas J.
Wild, Carl T.
Barney, Shawn O.
Lambert, Dennis M.
Petteaway, Stephen R.
Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TRANSMISSION
NUMBER OF SEQUENCES: 239
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/267,682
FILING DATE: 08-Oct-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/484,223A
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 167 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 109:
US-10-267-682-109
Query Match 94.8%; Score 110; DB 4; Length 167;
Best Local Similarity 91.7%; Pred. No. 7.4e-08;

Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAITIGILAAIALPAPYDY 24
 |||||:||||:|||||
 Db 11 IELMTVIAIVGILAAVALPAPYDY 34

RESULT 14

US-10-267-748-109
 ; Sequence 109, Application US/10267748
 ; Publication No. US20040052820A1

GENERAL INFORMATION:

APPLICANT: Bolognesi, Dani P.

Matchews, Thomas J.

Wild, Carl T.

Barney, Shawn O.

Lambert, Dennis M.

Petleway, Stephen R.

Langlois, Alphonse J.

MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV

TRANSMISSION

NUMBER OF SEQUENCES: 239

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/267,748

FILING DATE: 08-Oct-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/484,223A

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7872-029

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 109:

SEQUENCE CHARACTERISTICS:

LENGTH: 167 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: unknown

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 109:

US-10-267-748-109

Query Match 94.8%; Score 110; DB 4; Length 167;

Best Local Similarity 91.7%; Pred. No. 7.4e-08;

Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAITIGILAAIALPAPYDY 24
 |||||:||||:|||||

Db 11 IELMTVIAIVGILAAVALPAPYDY 34

RESULT 15

US-10-472-260-130

; Sequence 130, Application US/10472260

; Publication No. US20040265328A1

GENERAL INFORMATION:

APPLICANT: MICROBIOLOGICAL RESEARCH AUTHORITY

IMPERIAL COLLEGE INNOVATIONS LIMITED

TITLE OF INVENTION: IMMUNOGENIC COMMENSAL NEISSERIA SEQUENCES

FILE REFERENCE: GMS/DJC/23480

CURRENT APPLICATION NUMBER: US/10/472,260

CURRENT FILING DATE: 2003-09-22

NUMBER OF SEQ ID NOS: 199

SOFTWARE: Patentin version 3.1

SEQ ID NO 130

LENGTH: 170

TYPE: PRP

ORGANISM: Neisseria meningitidis (group B)

US-10-472-260-130

Query Match 94.8%; Score 110; DB 5; Length 170;

Best Local Similarity 91.7%; Pred. No. 7.6e-08;

Matches 22; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELMTVIAITIGILAAIALPAPYDY 24
 |||||:||||:|||||

Db 11 IELMTVIAIVGILAAVALPAPYDY 34

Search completed: January 23, 2006, 12:08:59
 Job time : 57.4236 secs

THE UNIVERSITY OF CHICAGO

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:53:39 / Search time 2.23561 Seconds
(without alignments)
72.521 Million cell updates/sec

Title: US-10-501-838a-15
Perfect score: 90
Sequence: 1 ASFGFCIGRLCVQDGF 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: /cgnt_6/prodata/2/pubppa/US08_NEW_PUB.pep.*
2: /cgnt_6/prodata/2/pubppa/US06_NEW_PUB.pep.*
3: /cgnt_6/prodata/2/pubppa/US07_NEW_PUB.pep.*
4: /cgnt_6/prodata/2/pubppa/PCT_NEW_PUB.pep.*
5: /cgnt_6/prodata/2/pubppa/US09_NEW_PUB.pep.*
6: /cgnt_6/prodata/2/pubppa/US10_NEW_PUB.pep.*
7: /cgnt_6/prodata/2/pubppa/US11_NEW_PUB.pep.*
8: /cgnt_6/prodata/2/pubppa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	43.3	474	7	US-11-043-889-16
2	39	43.3	1167	6	US-10-601-368-18
3	38	42.2	801	6	US-10-467-657-6470
4	38	42.2	1076	6	US-10-467-657-5708
5	37.5	41.7	167	7	US-11-108-172-204
6	37.5	41.7	310	6	US-10-878-556A-196
7	37.5	41.7	328	6	US-10-821-234-1462
8	37.5	41.7	328	6	US-10-878-556A-26
9	37.5	41.7	100	6	US-10-467-657-9162
10	37	41.1	276	7	US-11-062-186-40
11	37	41.1	277	7	US-11-033-039-858
12	37	41.1	1020	6	US-10-513-786-4
13	37	41.1	1900	6	US-10-513-786-3
14	37	41.1	3704	6	US-10-513-786-1
15	36	40.0	56	6	US-10-467-657-4332
16	36	40.0	141	7	US-11-084-508-22
17	36	40.0	434	6	US-10-467-657-3230
18	36	40.0	187	6	US-10-667-295-72
19	35	38.9	211	6	US-10-667-295-71
20	35	38.9	522	6	US-10-995-561-1030
21	35	38.9	733	6	US-10-821-234-1147
22	34	37.8	33	7	US-11-198-847-45
23	34	37.8	33	7	US-11-198-847-54
24	34	37.8	43	7	US-11-123-896-249
25	34	37.8	58	6	US-10-613-744-31

26	34	37.8	59	6	US-10-467-657-8062	Sequence 8062, App
27	34	37.8	109	7	US-11-123-896-248	Sequence 248, App
28	34	37.8	145	7	US-11-084-508-21	Sequence 51, App1
29	34	37.8	235	7	US-11-219-995-5	Sequence 5, App1
30	34	37.8	260	6	US-10-485-517-354	Sequence 354, App
31	34	37.8	261	6	US-10-485-517-150	Sequence 150, App
32	34	37.8	376	7	US-11-109-157A-12	Sequence 12, App1
33	34	37.8	481	6	US-10-467-657-3916	Sequence 3916, App
34	34	37.8	495	7	US-11-219-995-2	Sequence 2, App1
35	34	37.8	556	6	US-10-613-744-8	Sequence 8, App1
36	34	37.8	608	7	US-11-109-157A-11	Sequence 11, App1
37	34	37.8	619	7	US-11-109-157A-42	Sequence 42, App1
38	34	37.8	871	7	US-11-109-157A-10	Sequence 10, App1
39	34	37.8	1044	7	US-11-091-668-2	Sequence 2, App1
40	34	37.8	1103	7	US-11-109-157A-9	Sequence 9, App1
41	34	37.8	1250	7	US-11-137-465-62	Sequence 62, App1
42	34	37.8	1313	7	US-11-091-668-4	Sequence 4, App1
43	33.5	37.2	48	7	US-11-123-896-282	Sequence 282, App
44	33.5	37.2	76	7	US-11-123-896-281	Sequence 281, App
45	33.5	37.2	707	7	US-11-186-284-132	Sequence 132, App

ALIGNMENTS

RESULT 1
US-11-043-889-16
Sequence 16, Application US/11043889
Publication No. US2006000819A1
GENERAL INFORMATION:
APPLICANT: Currie, Rory A.J.
APPLICANT: Gluckman, Maria Alexandra
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
TITLE OF INVENTION: 57259, 67118, 67067, 62092, FBH58295FL, 57255,
TITLE OF INVENTION: AND 57255a1 MOLECULES AND USES THEREFOR
FILE REFERENCE: MP102-095DVIOWNIM
CURRENT APPLICATION NUMBER: US/11/043, 889
CURRENT FILING DATE: 2005-01-26
PRIOR APPLICATION NUMBER: US 10/154, 419
PRIOR FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 09/858194
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: US 60/204211
PRIOR FILING DATE: 2000-05-12
PRIOR APPLICATION NUMBER: US 09/895811
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215376
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 09/919781
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/221769
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 09/957664
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: US 60/233790
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: US 09/964295
PRIOR FILING DATE: 2001-09-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 59
SOFTWARE: FastSeq Version 4.0
SEQ ID NO 16
LENGTH: 474
TYPE: PRT
ORGANISM: Homo sapiens
US-11-043-889-16
Query Match 43.3%; Score 39; DB 7; Length 474;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
CY 2 SPFGICIGRLCVQD 14

Db 385 SFSSTAGRCITBD 397

RESULT 2
US-10-601-368-18
; Sequence 18, Application US/10601368
; Publication No. US20050260702A1
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; APPLICANT: Lora, Jose M.
; TITLE OF INVENTION: NOVEL INTEGRIN ALPHA SUBUNIT AND USES THEREOF
; FILE REFERENCE: 07334-275001
; CURRENT APPLICATION NUMBER: US/10/601,368
; PRIOR FILING DATE: 2003-06-23
; PRIOR APPLICATION NUMBER: US/09/561,263A
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/322,790
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-601-368-18

Query Match 43.3%; Score 39; DB 6; Length 1167;
Best Local Similarity 55.6%; Pred. No. 65;
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 1 ASFGFCIGRL--CVDDGF 16
| | | | |
Db 547 ARFGFAMGALPDLNDGF 564

RESULT 3
US-10-467-657-6470
; Sequence 6470, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MONACI Elisabetta
; APPLICANT: MASIGNANI Vega
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6470
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6470

Query Match 42.2%; Score 38; DB 6; Length 801;
Best Local Similarity 46.2%; Pred. No. 67;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 SFGFCIGRLCVD 14
: | | | | |
Db 13 ALGFCFGTHCAAD 25

RESULT 4
US-10-467-657-5708
; Sequence 5708, Application US/10467657
; Publication No. US20050260581A1

; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MONACI Elisabetta
; APPLICANT: MASIGNANI Vega
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5708
; LENGTH: 1076
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5708

Query Match 42.2%; Score 38; DB 6; Length 1076;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 FGFCIGR 9
| | | | |
Db 802 FGFCGR 808

RESULT 5
US-11-108-172-204
; Sequence 204, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon B.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLORECTAL CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10

```
/ PRIOR APPLICATION NUMBER: US 09/476,296
/ PRIOR FILING DATE: 1999-12-30
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1130
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 204
/ LENGTH: 167
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-821-234-1462

Query Match          41.7%; Score 37.5; DB 6; Length 167;
Best Local Similarity 37.0%; Pred. No. 21;
Matches 10; Conservative 0; Mismatches 6; Indels 11; Gaps 1;

QY 1 ASFGFCIG-----RLCVQDGF 16
   |||||
Db 121 AVHGCCTGGGVDLVTACDIRYCAQDAF 147

RESULT 6
US-10-878-556A-196
/ Sequence 196, Application US/10878556A
/ Publication No. US2005026639A1
/ GENERAL INFORMATION:
/ APPLICANT: Hoffmann La-Roche Inc.
/ TITLE OF INVENTION: HCV regulated protein expression
/ FILE REFERENCE: 21762
/ CURRENT APPLICATION NUMBER: US/10/878,556A
/ CURRENT FILING DATE: 2004-06-28
/ NUMBER OF SEQ ID NOS: 199
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 196
/ LENGTH: 310
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: humangp/chri9-q8wx0
/ DATABASE ENTRY DATE: 2003-04-22
US-10-878-556A-196

Query Match          41.7%; Score 37.5; DB 6; Length 310;
Best Local Similarity 37.0%; Pred. No. 35;
Matches 10; Conservative 0; Mismatches 6; Indels 11; Gaps 1;

QY 1 ASFGFCIG-----RLCVQDGF 16
   |||||
Db 148 AVHGCCTGGGVDLVTACDIRYCAQDAF 174

RESULT 7
US-10-821-234-1462
/ Sequence 1462, Application US/10821234
/ Publication No. US2005025511A1
/ GENERAL INFORMATION:
/ APPLICANT: Labat, Ivan
/ APPLICANT: Steach-Crain, Birgit
/ APPLICANT: Andarmant, Susan
/ APPLICANT: Tang, Y. Tom
/ TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
/ FILE REFERENCE: 821A
/ CURRENT APPLICATION NUMBER: US/10/821,234
/ CURRENT FILING DATE: 2004-04-07
/ PRIOR APPLICATION NUMBER: US 60/462,047
/ PRIOR FILING DATE: 2003-04-07
/ NUMBER OF SEQ ID NOS: 1704
/ SOFTWARE: PC_SEQ_genes Version 1.0
/ SEQ ID NO 1462
/ LENGTH: 328
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-821-234-1462
```

```
Query Match          41.7%; Score 37.5; DB 6; Length 328;
Best Local Similarity 37.0%; Pred. No. 37;
Matches 10; Conservative 0; Mismatches 6; Indels 11; Gaps 1;

QY 1 ASFGFCIG-----RLCVQDGF 16
   |||||
Db 166 AVHGCCTGGGVDLVTACDIRYCAQDAF 192

RESULT 8
US-10-878-556A-26
/ Sequence 26, Application US/10878556A
/ Publication No. US2005026639A1
/ GENERAL INFORMATION:
/ APPLICANT: Hoffmann La-Roche Inc.
/ TITLE OF INVENTION: HCV regulated protein expression
/ FILE REFERENCE: 21762
/ CURRENT APPLICATION NUMBER: US/10/878,556A
/ CURRENT FILING DATE: 2004-06-28
/ NUMBER OF SEQ ID NOS: 199
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 26
/ LENGTH: 328
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: sw_hum/ech1_human
/ DATABASE ENTRY DATE: 1997-11-01
US-10-878-556A-26

Query Match          41.7%; Score 37.5; DB 6; Length 328;
Best Local Similarity 37.0%; Pred. No. 37;
Matches 10; Conservative 0; Mismatches 6; Indels 11; Gaps 1;

QY 1 ASFGFCIG-----RLCVQDGF 16
   |||||
Db 166 AVHGCCTGGGVDLVTACDIRYCAQDAF 192

RESULT 9
US-10-467-657-9162
/ Sequence 9162, Application US/10467657
/ Publication No. US20050260581A1
/ GENERAL INFORMATION:
/ APPLICANT: CHIRON SPA
/ APPLICANT: FONTANA Maria Rita
/ APPLICANT: PIZZA Mariagrazia
/ APPLICANT: MASIGNANI Vega
/ APPLICANT: MONACI Elisabetta
/ TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/467,657
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ PRIOR FILING DATE: 2001-02-12
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqWin99, version 1.04
/ SEQ ID NO 9162
/ LENGTH: 100
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-467-657-9162

Query Match          41.1%; Score 37; DB 6; Length 100;
Best Local Similarity 36.4%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 3; Indels 8; Gaps 1;

QY 3 FGR-----CIGRLCVQDGF 16
   |||||
Db 59 FGRFRRPTAIPRLDRICIEQGF 80

RESULT 10
```

US-11-062-186-40
; Sequence 40, Application US/11062186
; Publication No. US20050272097A1
; GENERAL INFORMATION:
; APPLICANT: CALENOFF, EMANUEL
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING AND TREATING
; TITLE OF INVENTION: AUTOIMMUNE DISEASES
; FILE REFERENCE: 21417-98470
; CURRENT APPLICATION NUMBER: US/11/062,186
; CURRENT FILING DATE: 2005-02-18
; PRIOR APPLICATION NUMBER: 60/546,062
; PRIOR FILING DATE: 2004-02-18
; PRIOR APPLICATION NUMBER: 60/545,980
; PRIOR FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 40
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-062-186-40

Query Match 41.1%; Score 37; DB 7; Length 276;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ASFGFCIGRLCYQDGF 16
||| | | | : | | |
DB 75 ASFFFLYGALLAEGF 90

RESULT 11
US-11-033-039-858
; Sequence 858, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPIOTOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 858
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-033-039-858

Query Match 41.1%; Score 37; DB 7; Length 277;
Best Local Similarity 50.0%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ASFGFCIGRLCYQDGF 16
||| | | | : | | |
DB 76 ASFFFLYGALLAEGF 91

RESULT 12
US-10-513-786-4
; Sequence 4, Application US/10513786
; Publication No. US20050260589A1
; GENERAL INFORMATION:
; APPLICANT: Bio Control Institute Limited
; APPLICANT: OHTA, Michio
; APPLICANT: AGATA, Norio

; TITLE OF INVENTION: A cereulide synthetase, a gene thereof, and a detection method fo
; TITLE OF INVENTION: cereulide.
; FILE REFERENCE: P0202401
; CURRENT APPLICATION NUMBER: US/10/513,786
; CURRENT FILING DATE: 2004-11-16
; PRIOR APPLICATION NUMBER: JP P2002-142398
; PRIOR FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 4
; LENGTH: 1020
; TYPE: PRT
; ORGANISM: Bacillus cereus
US-10-513-786-4

Query Match 41.1%; Score 37; DB 6; Length 1020;
Best Local Similarity 45.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 CIGRLCYQDGF 16
| : | : | : | : | :
DB 336 CLGGICVAKGY 346

RESULT 13
US-10-513-786-3
; Sequence 3, Application US/10513786
; Publication No. US20050260589A1
; GENERAL INFORMATION:
; APPLICANT: Bio Control Institute Limited
; APPLICANT: OHTA, Michio
; APPLICANT: AGATA, Norio
; TITLE OF INVENTION: A cereulide synthetase, a gene thereof, and a detection method fo
; TITLE OF INVENTION: cereulide.
; FILE REFERENCE: P0202401
; CURRENT APPLICATION NUMBER: US/10/513,786
; CURRENT FILING DATE: 2004-11-16
; PRIOR APPLICATION NUMBER: JP P2002-142398
; PRIOR FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 3
; LENGTH: 1900
; TYPE: PRT
; ORGANISM: Bacillus cereus
US-10-513-786-3

Query Match 41.1%; Score 37; DB 6; Length 1900;
Best Local Similarity 45.5%; Pred. No. 2e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 CIGRLCYQDGF 16
| : | : | : | : | :
DB 336 CLGGICVAKGY 346

RESULT 14
US-10-513-786-1
; Sequence 1, Application US/10513786
; Publication No. US20050260589A1
; GENERAL INFORMATION:
; APPLICANT: Bio Control Institute Limited
; APPLICANT: OHTA, Michio
; APPLICANT: AGATA, Norio
; TITLE OF INVENTION: A cereulide synthetase, a gene thereof, and a detection method fo
; TITLE OF INVENTION: cereulide.
; FILE REFERENCE: P0202401
; CURRENT APPLICATION NUMBER: US/10/513,786
; CURRENT FILING DATE: 2004-11-16
; PRIOR APPLICATION NUMBER: JP P2002-142398
; PRIOR FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1

SEQ ID NO 1
LENGTH: 3704
TYPE: PRT
ORGANISM: Bacillus cereus
US-10-513-786-1

Query Match 41.1%; Score 37; DB 6; Length 3704;
Best Local Similarity 45.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 CIGRLCVQDGF 16
|:|:|:|:|:
DB 2140 CIGGICVAKGY 2150

RESULT 15
US-10-467-657-4332
Sequence 4332, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 4332
LENGTH: 56
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4332

Query Match 40.0%; Score 36; DB 6; Length 56;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 FCIGRLCVQDGF 16
|:|:|:|:|:
DB 26 FSSGLFCVWMDGF 37

Search completed: January 23, 2006, 12:09:42
Job time : 3.23581 secs

THE '80s Blank (upstg)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:20:59 ; Search time 53.1004 Seconds
(without alignments)
132.392 Million cell updates/sec

Title: US-10-501-838a-15
Perfect score: 90
Sequence: 1 ASFGPCIGRLCVQDGF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*
9: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	16	ADBI6900	AdBI6900 Zonula oc
2	90	100.0	16	AEBO8232	AeBO8232 Zonula oc
3	90	100.0	383	ADW9312	AdW9312 CTX-phi 2
4	90	100.0	399	AAK20006	AaK20006 Zonula oc
5	70	77.8	14	AAOI4100	AaOI4100 N-termina
6	70	77.8	12	ADW9317	AdW9317 Delta-G I
7	50	55.6	53	AAO09038	AaO09038 Human pol
8	48	53.3	238	ABU3427	AbU3427 Protein e
9	48	53.3	238	AEBA0700	AeBA0700 L. pneumo
10	48	53.3	238	AEBA0700	AeBA0700 L. pneumo
11	45	50.0	8	AAV79110	AaV79110 Peptide a
12	45	50.0	8	AAOI4073	AaOI4073 Human zon
13	45	50.0	240	ABO77295	AbO77295 Pseudomon
14	45	50.0	751	AAAG39904	AaAG39904 Arabidops
15	45	50.0	781	AAAG39903	AaAG39903 Arabidops
16	45	50.0	837	AAAG39902	AaAG39902 Arabidops
17	44	48.9	131	ABG76547	AbG76547 HCV E1 an
18	44	48.9	213	ADY23274	AdY23274 Plant ful
19	44	48.9	238	ABU33372	AbU33372 Protein e
20	44	48.9	238	ABU33372	AbU33372 L. pneumo
21	44	48.9	243	ABU37023	AbU37023 L. pneumo
22	44	48.9	492	ADN46868	AdN46868 Thermococ
23	44	48.9	525	ADM06103	AdM06103 Human pro
24	43	47.8	104	ABP00903	AbP00903 Human ORF

25	43	47.8	284	3	AAG24558
26	43	47.8	421	7	ADM25652
27	43	47.8	432	7	AAU35671
28	43	47.8	432	6	ABU39213
29	43	47.8	432	6	ABU30641
30	43	47.8	484	6	ABR52804
31	43	47.8	484	7	ADK62196
32	43	47.8	520	5	ABP74028
33	43	47.8	944	7	ADK08076
34	43	47.8	944	7	ADU40446
35	42.5	47.2	84	2	AAV30424
36	42.5	47.2	162	2	AAV30437
37	42.5	47.2	162	3	AAV35322
38	42.5	47.2	181	2	AAV31712
39	42.5	47.2	181	2	AAV30410
40	42	46.7	8	3	AAV79106
41	42	46.7	8	5	AAOI4069
42	42	46.7	63	9	ADY85736
43	42	46.7	76	8	ADY06638
44	42	46.7	85	9	ADY85734
45	42	46.7	310	7	ADC79374

ALIGNMENTS

RESULT 1
ADBI6900 standard; peptide; 16 AA.
XX
XX
ADBI6900;
XX

20-NOV-2003 (first entry)
XX
XX
Zonula occludens toxin (ZOT) penetrating peptide 15.
XX

penetrating peptide; epithelial; endothelial; tight junction; diabetes;
XX
XX
infertility; hormone; vitamin deficiency; neurodegenerative;
XX
XX
cardiovascular; haematological; endocrine disorder; obesity;
XX
XX
neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
XX
XX
osteopathic; cytoskeletal; nootropic.

OS unidentified.

PN WO2003066859-A2.

PD 14-AUG-2003.

PF 07-FEB-2003; 2003WO-IB000966.

PR 07-FEB-2002; 2002US-0355396P.

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PI Ben-Sasson SA, Cohen E;

DR WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating
or preventing e.g. endocrine disorders.

PS Claim 2; Page 14; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
translocating across a biological barrier. Furthermore, it refers to
methods that use these peptides to facilitate penetration of a
biologically active effector molecule such as a drug or other therapeutic
agent across biological barriers e.g. epithelial or endothelial cells
sealed by tight junctions. This peptide is derived from a bacterial
toxin, an integral membrane or extracellular protein and can comprise an
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
or enzyme. The effector molecule, however, can comprise for example
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
CC cyostatic or neurotropic activities. This peptide is from the zonula
CC occludens toxin and is penetrating peptide 15 of the invention.
CC
CC Sequence 16 AA:
SQ
Query Match 100.0%; Score 90; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ASFGFCIGRLCYVDGF 16
Db 1 ASFGFCIGRLCYVDGF 16
RESULT 2
AEB08232
ID AEB08232 standard; peptide; 16 AA.
XX
AC AEB08232;
XX
DT 25-AUG-2005 (first entry)
XX
DE Zonula occludens toxin (ZOT) penetrating peptide 15, SEQ ID NO: 15.
XX
KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
KW antidiabetic; endocrine disease; gastrointestinal disease;
KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
KW neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
KW anticonvulsant; genetic disorder; cardiovascular disease;
KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
KW genitourinary disease; hematological disease; antianemic; anemia;
KW autoimmune disease; immunosuppressive; immune deficiency;
KW immunostimulant; infectious disease; antimicrobial; infection;
KW erectile dysfunction; andrology; major depressive disorder;
KW antidepressant; psychiatric disorder; pain; analgesic;
KW bacterial infection; antibacterial; viral infection; virocidic;
KW fungal infection; fungicide; parasitic infection; antiparasitic;
KW renal failure; antifertility; antipneumatic; cyostatic;
KW antiinflammatory; hepatotropic; hepatitis B virus infection.
XX
OS Vibrio cholerae.
XX
PN US2005136103-A1.
XX
PD 23-JUN-2005.
XX
PF 16-SEP-2004; 2004US-00942300.
XX
PR 17-SEP-2003; 2003US-00664989.
XX
PR 17-SEP-2003; 2003US-00665184.
XX
PR 17-SEP-2003; 2003US-0503615P.
XX
PA (BENS/) BEN-SASSON S A.
XX
PA (COHE/) COHEN E.
XX
PI Ben-Sasson SA, Cohen E;
XX
XX WPI; 2005-444089/45.
XX
XX Composition used for translocating effectors across barrier such as
XX epithelial cells during treatment of e.g. endocrine disorders comprises
XX

PT effector sequentially coupled with counter ion and hydrophobic agent.
XX
XX Claim 46; SEQ ID NO 15; 59pp; English.
XX
XX The present invention relates to a pharmaceutical composition of
XX penetrating peptides for transseptal delivery of effector. The
XX invention comprises the effector sequentially coupled with a counter ion
XX and at least one hydrophobic agent, where the effector is selectively
XX encapsulated into a complex. The invention is useful for translocating
XX effectors across a biological barrier such as epithelial cells and
XX endothelial cells during treatment and prevention of disease or
XX pathological conditions (including endocrine disorders, diabetes,
XX infertility, hormone deficiencies, osteoporosis, ophthalmological
XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
XX Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
XX coagulable states, coronary disease, cerebrovascular events, metabolic
XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
XX hematological disorders, anemia of different entities, immunologic and
XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
XX infectious diseases, viral infections, bacterial infections, fungal
XX infections, parasitic infections, neoplastic diseases, multi-factorial
XX disorders, impotence, chronic pain, depression, different fibrosis states
XX and short stature) and for mucosal vaccination against anthrax and
XX hepatitis B. The present sequence is the Vibrio cholerae zonula occludens
XX toxin (ZOT) penetrating peptide. This sequence is used in the effective
XX translocation of aminoglycoside antibiotics and antifungal agents across
XX an epithelial barrier.
XX
XX Sequence 16 AA:
SQ
Query Match 100.0%; Score 90; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ASFGFCIGRLCYVDGF 16
Db 1 ASFGFCIGRLCYVDGF 16
RESULT 3
ADW99312
ID ADW99312 standard; protein; 383 AA.
XX
AC ADW99312;
XX
DT 21-APR-2005 (first entry)
XX
DE CTR-phl ZOT protein.
XX
XX Cytostatic; Immunosuppressive; Vasotropic; Antibacterial;
KW Antiinflammatory; Gastrointestinal-gen.; dermatological; Antiallergic;
KW Antiasthmatic; Gene therapy; Vaccine; zonulin receptor;
KW CTR-phl ZOT protein; food allergy; gastrointestinal infection;
KW inflammation; autoimmune disease; inflammatory bowel disease;
KW Celiac Disease; systemic lupus erythematosus; asthma;
KW irritable bowel syndrome.
XX
OS Vibrio cholerae.
XX
PN WO2005010022-A2.
XX
PD 03-FEB-2005.
XX
PF 15-JUL-2004; 2004WO-US022753.
XX
PR 15-JUL-2003; 2003US-0487889P.
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
XX Paaano A, Vogel SN;
XX
XX WPI; 2005-123116/13.
XX

XX New agonist polypeptide of a human receptor of zonulin and Vibrio
PT cholerae phage CTXph ZOT protein, useful for diagnosing, preventing or
PT treating diseases associated with increased expression of zonulin, e.g.
PT autoimmune diseases.

PS Claim 3, SEQ ID NO 1, 55bp, English.

XX This sequence represents Vibrio cholerae phage CTX-ph ZOT protein. The
CC agonist polypeptide of the invention is less than 100 amino acid residues
CC in length and does not contain residues 294-298 of this protein. A
CC pharmaceutical composition for treating a disease, or a vaccine, comprising
CC a therapeutic agent for treating the disease, and the agonist
CC polypeptide of the invention. The composition of the invention may be
CC used in the treatment of a food allergy, a gastrointestinal infection or
CC inflammation, an autoimmune disease, inflammatory bowel disease or Celiac
CC disease. The therapeutic agent is a drug, a biological peptide, an anti-
CC antibody, an antibody fragment, a single chain antibody (scFv), an anti-
CC cancer drug, an antibiotic, a hormone or a cytokine. The composition and
CC methods of the invention are useful for diagnosing, preventing or
CC treating diseases or disorders associated with increased expression of
CC zonulin, such as cancer, autoimmune diseases, vascular diseases, Celiac
CC bacterial infection, gastritis, inflammatory bowel disease, Celiac
CC disease, systemic lupus erythematosus, food allergy, asthma or irritable
CC bowel syndrome. These may also be used in drug screening purposes.

XX Sequence 363 AA;

Query Match 100.0%; Score 90; DB 9; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASFGFCIGRLCVQDGF 16
DB 268 ASFGFCIGRLCVQDGF 283

RESULT 1
AAR20006
ID AAR20006 standard; protein; 399 AA.

XX AAR20006;
XX 25-MAR-2003 (revised)
XX 31-MAR-1992 (first entry)
XX Zonula occludens toxin.

XX ZOT; cholera; vaccine; enterotoxin; diarrhoea.
XX Vibrio cholerae.
XX WO918979-A.

XX 12-DEC-1991.

XX 05-JUN-1990; 90US-00533315.

XX 05-JUN-1990; 90US-00533315.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Kaper JB, Baudrymaur B, Fasano A;

XX WPI, 1992-007465/01.

XX N-PSDB; AAQ20185.

XX New Vibrio cholerae strains - comprise restriction endonuclease fragment
PT encoding toxin, used as vaccines against cholera.

XX Disclosure, Fig 18, 83pp; English.

XX The amino acid sequence is that of the zonula occludens toxin (ZOT). It

CC may be responsible for diarrhoea in some strains of cholera and the ZOT
CC gene or fragments of it are deleted from strains of Vibrio cholerae
CC (V.c.) to be used as vaccines. These V.c. strains have 100% efficacy in
CC protecting humans against subsequent infection with a strain of a similar
CC serotype and avoid undesirable side effects such as diarrhoea, nausea and
CC cramping. Cultures of these strains may be used for prodn. of vaccines
CC against cholera. (updated on 25-MAR-2003 to correct PI field.)

XX Sequence 399 AA;

Query Match 100.0%; Score 90; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASFGFCIGRLCVQDGF 16
DB 284 ASFGFCIGRLCVQDGF 299

RESULT 5
AA014100
ID AA014100 standard; peptide; 14 AA.

XX AA014100;
XX 07-MAY-2002 (first entry)
XX N-terminal region of zonulin from V. cholerae ZOT.

XX Zonulin; synthetic octapeptide; antagonist; zonula occludens toxin; ZOT;
XX mammalian tight junction; rat; paracellular permeability; jejunum;
XX type I diabetes; Vibrio cholerae.

XX Vibrio cholerae.
XX WO200189551-A1.

XX 29-NOV-2001.

XX 21-MAR-2001; 2001WO-US005825.

XX 19-MAY-2000; 2000US-0205305P.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Fasano A, Watts TL;

XX WPI, 2002-062480/08.

XX Prevention or delay of onset of diabetes, preferably type I diabetes,
XX comprises administering to a subject a peptide antagonist of zonulin
XX which binds to zonula occludens toxin receptor.

XX Disclosure, Fig 1, 47pp; English.

XX This sequence represents a synthetic peptide from the N-terminal region
CC of zonulin from Vibrio cholerae ZOT. The invention relates to preventing
CC or delaying the onset of diabetes, comprising administering to a subject
CC a peptide antagonist of zonulin which binds to zonula occludens toxin
CC (ZOT) receptor. The peptide antagonist of zonulin of the invention, which
CC competitively inhibits the binding of ZOT and zonulin to the ZOT
CC receptor, inhibits the ability of ZOT and zonulin to physiologically
CC modulate the opening of mammalian tight junctions (TJ). By using diabetic
CC -prone rats and treating them with the peptide antagonist of zonulin of
CC the invention, it was shown there is a significant increase in
CC paracellular permeability of the small intestine restricted to the
CC jejunum. The peptide antagonist was able to effectively block the
CC permeability changes required for the development of diabetes. The method
CC is useful for preventing or delaying the onset of type I diabetes

XX Sequence 14 AA;

Query Match 77.8%; Score 70; DB 5; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FCIGRLCVQDGF 16
Db 1 FCIGRLCVQDGF 12

RESULT 6
ADM99317
ID ADM99317 standard; peptide; 112 AA.

AC ADM99317;

DT 21-APR-2005 (first entry)

DE Delta-G fragment of ZOT protein.

KW Cytostatic; Immunosuppressive; Vasoactive; Antibacterial;

KW Antiinflammatory; Gastrointestinal-Gen.; Dermatological; Allergic;

KW Antiasthmatic; Gene therapy; Vaccine; zonulin receptor;

KW Cyt-pm ZOT protein; food allergy; gastrointestinal infection;

KW inflammation; autoimmune disease; inflammatory bowel disease;

KW Celiac Disease; systemic lupus erythematosus; asthma;

KW Irritable bowel syndrome.

XX Vibrio cholerae.

XX WO2005010022-A2.

XX 03-FEB-2005.

XX 15-JUL-2004; 2004WO-US022753.

XX 15-JUL-2003; 2003US-0487889P.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Fasano A, Vogel SN;

XX WPI; 2005-123116/13.

XX Claim 58; SEQ ID NO 6; 55bp; English.

XX This sequence represents the delta-G fragment of Vibrio cholerae phage

XX ZOT protein. This protein fragment was used in a method of the invention

XX for screening for drug candidates for treating a disease. A

XX pharmaceutical composition for treating a disease, and a vaccine,

XX comprises a therapeutic agent for treating the disease, and the agonist

XX polypeptide of the invention. The composition of the invention may be

XX used in the treatment of a food allergy, a gastrointestinal infection or

XX inflammation, an autoimmune disease, inflammatory bowel disease or Celiac

XX Disease. The therapeutic agent is a drug, a biological peptide, an

XX antibody, an antibody fragment, a single chain antibody (ScFv), an anti-

XX cancer drug, an antibiotic, a hormone or a cytokine. The composition and

XX methods of the invention are useful for diagnosing, preventing or

XX treating diseases or disorders associated with increased expression of

XX zonulin, such as cancer, autoimmune diseases, vascular diseases,

CC bacterial infection, gastritis, inflammatory bowel disease, Celiac

CC Disease, systemic lupus erythematosus, food allergy, asthma or irritable

CC bowel syndrome. These may also be used in drug screening purposes.

CC Sequence 112 AA;

CC Query Match 77.8%; Score 70; DB 9; Length 112;

CC Best Local Similarity 100.0%; Pred. No. 0.0052;

CC Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 FCIGRLCVQDGF 16
Db 1 FCIGRLCVQDGF 12

RESULT 7
AAO09038
ID AAO09038 standard; protein; 53 AA.

AC AAO09038;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 22930.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukemia;

KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AA188969.

XX Claim 20; SEQ ID NO 22930; 1399pp + Sequence listing; English.

XX The invention relates to human polynucleotides (AA179941-AA19341) and

XX the encoded proteins (AA000010-AA013910) that exhibit activity relating to

XX cytokine, cell proliferation or cell differentiation or which may induce

XX production of other cytokines in other cell populations. The

XX polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX peptide therapy. The polypeptides have various cytokine-like activities,

XX e.g. stem cell growth factor activity, hematopoietic regulatory

XX activity, tissue growth factor activity, immunomodulatory activity and

XX activin/inhibin activity and may be useful in the diagnosis and/or

XX treatment of cancer, leukemia, nervous system disorders, arthritis and

XX inflammation. Note: The sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic format

XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 53 AA;

XX Query Match 55.6%; Score 50; DB 4; Length 53;

XX Best Local Similarity 52.6%; Pred. No. 3.2;

XX Matches 10; Conservative 2; Mismatches 3; Indels 4; Gaps 1;

OY 1 ASFGCTGRGC---VQDG 15

Db 2 ASFGCTGRGC---VQDG 15

RESULT 8

ID ABU33427 standard; protein; 238 AA.

AC ABU33427;

XX 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #18954.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Legionella pneumophila.
 XX WO200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 XX 25-OCT-2001; 2001US-0342923P.
 XX 08-FEB-2002; 2002US-00072851.
 XX 06-MAR-2002; 2002US-0362699P.
 XX (BLT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KJ, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX MPI; 2003-029926/02.
 XX N-PSDB; ACA37287.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 61351; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 238 AA;

Query Match 53.3%; Score 48; DB 6; Length 238;
 Best Local Similarity 57.1%; Pred. No. 26;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ASFGRCIGRLCYVD 14
 ID AEB40700 standard; protein; 238 AA.
 DB 115 AATGTCFSGSLCYLD 128

RESULT 9

XX AEB40700
 XX ID AEB40700 standard; protein; 238 AA.
 XX AC AEB40700;
 XX DT 08-SEP-2005 (first entry)
 XX DE L. pneumophila protein SEQ ID NO 5032.
 XX detection; infection; Antibacterial; Vaccine.
 XX Legionella pneumophila.
 XX OS WO2005049642-A2.
 XX PN WO2005049642-A2.
 XX 02-JUN-2005.
 XX PD 23-SEP-2004; 2004WO-1B003578.
 XX PR 21-NOV-2003; 2003FR-00013687.
 XX (INSP) INST PASTEUR.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX Buchrieser C, Ticht M, Etienne J, Ma L, Cazalet C, Glaeser P,
 PI Rusniok C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F,
 PI Jarraud S;
 XX MPI; 2005-388305/40.
 XX New genome of Legionella pneumophila Paris strain and derived
 PT polypeptides, useful for detection or identification of the strain and
 PT for treatment and prevention of infections.
 XX Claim 3; SEQ ID NO 5032; 660pp; English.
 XX The invention relates to an isolated or purified nucleotide sequences (1)
 CC from Legionella pneumophila Paris strain. (1), and their related
 CC sequences or fragments, are useful as primers and probes for detection
 CC and amplification, including differentiation between the Paris and
 CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
 CC (hybrid) polypeptides (11). (11) are also useful for preparation of
 CC specific antibodies (Ab), also used for detection/identification of
 CC Legionella, and some (11), specifically those involved in synthesis of
 CC surface proteins, are targets for identification of inhibitors. (11), or
 CC vectors that contain (11), are useful as vaccines and immunogenic
 CC compositions, for treatment and prevention of infections by L.
 CC pneumophila. The present sequence represents the amino acid sequence of a
 CC L. pneumophila protein.
 XX Sequence 238 AA;

Query Match 53.3%; Score 48; DB 9; Length 238;
 Best Local Similarity 57.1%; Pred. No. 26;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ASFGRCIGRLCYVD 14
 DB 115 AATGTCFSGSLCYLD 128

RESULT 10
 AEB37382
 ID AEB37382 standard; protein; 238 AA.

AC AEB37382;
XX 08-SEP-2005 (first entry)
XX L. pneumophila protein SEQ ID NO 1714.
DE detection; infection; Antibacterial; Vaccine.
XX Legionella pneumophila.
XX WO2005049642-A2.
XX 02-JUN-2005.
XX 23-SEP-2004; 2004WO-1B003578.
XX 21-NOV-2003; 2003FR-00013687.
XX (INSP) INST PASTEUR.
PA (INRM) INSEEM INST NAT SANTE & RECH MEDICALE.
PA (UJLY-) UNIV LYON 1 BERNARD CLAUDE.
XX (CNRS) CNRS CENT NAT RECH SCT.
PI Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P,
PI Ruenick C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F,
PI Jarraud S;
XX WPI; 2005-388305/40.
XX New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX Claim 3; SEQ ID NO 1714; 660bp; English.
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (I), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents the amino acid sequence of a
CC L. pneumophila protein.
XX
SQ Sequence 238 AA;
Query Match 53.3%; Score 48; DB 9; Length 238;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 ASFGFCIGRLCYOP 14
| : : : : :
DB 115 AAIGTCFGLCYLD 128
RESULT 11
AA79110
ID AA79110 standard; peptide; 8 AA.
XX
AC AA79110;
XX
DT 05-JUN-2000 (first entry)
XX
DE Peptide antagonist of zonulin.
XX Zonulin; antagonist; zonula occludens toxin receptor;
KW blood-brain barrier; antiinflammatory; cerebroprotective;
KW neuroprotective; dermatological; antiulcer; antiviral; antibacterial;

KW cyostatic; anti-HIV; vulnerary; antiallergic; hypotensive;
KW immunosuppressive; antiparasitic; vasotropic;
KW gastrointestinal inflammation; therapy.
XX Synthetic.
XX WO200007609-A1.
XX 17-FEB-2000.
XX 28-JUL-1999; 99WO-US016683.
XX 03-AUG-1998; 98US-00127815.
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX Fasano A;
XX WPI; 2000-205565/18.
XX
XX New peptide antagonist of zonulin useful as antiinflammatory agent for
PT treating cerebral ischemia, stroke, cerebral edema, gastritis,
PT shigellosis, viral gastroenteritis, meningitis, encephalomyelitis.
XX Claim 1; Page 42; 69pp; English.
XX This present sequence is that of a peptide antagonist of zonulin (Z), one
CC of 25 such peptides (see AA79105-29) of the invention, which bind to a
CC zonula occludens toxin (ZOT) receptor, yet do not physiologically
CC modulate the opening of mammalian tight junctions (TJ). The peptide
CC antagonists are based on a common motif of ZOT and human zonulins, which
CC is believed to be critical for receptor binding. They can be prepared by
CC chemical synthesis or by use of recombinant DNA techniques. The peptide
CC antagonists are used as an antiinflammatory agents in the treatment of
CC gastrointestinal inflammation, where they bind to the ZOT receptor in the
CC intestine and yet does not physiologically modulate the opening of TJ in
CC the intestine. Gastrointestinal inflammation conditions give rise to
CC increased intestinal permeability and the peptide is useful for treating
CC intestinal conditions that cause protein losing enteropathy caused by
CC infection, e.g. Clostridium difficile infection, enterocolitis,
CC shigellosis, viral gastroenteritis, parasite infestation, bacterial
CC overgrowth, whipple's disease, diseases with mucosal erosion or
CC ulcerations, e.g. gastritis, gastric cancer, collagenous colitis,
CC inflammatory bowel disease, diseases marked by lymphatic obstruction,
CC e.g. congenital intestinal lymphangiectasia, sarcoidosis lymphoma,
CC meenteric tuberculosis, and after surgical correction of congenital
CC heart disease with Fontan's operation, mucosal diseases without
CC ulceration, e.g. Menetrier's disease, coeliac disease, eosinophilic
CC gastroenteritis, and immune diseases, e.g. systemic lupus erythematosus
CC or food allergies, primarily to milk
XX
SQ Sequence 8 AA;
Query Match 50.0%; Score 45; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 GRLCYOPG 15
| : : : : :
DB 1 GRLCYOPG 8
RESULT 12
AA014073
ID AA014073 standard; peptide; 8 AA.
XX
AC AA014073;
XX
DT 07-MAY-2002 (first entry)
XX
DE Human zonulin peptide antagonist #6.
XX Zonulin; synthetic octapeptide; antagonist; zonula occludens toxin; ZOT;

PR 16-APR-1999;	99US-0129845P.
PR 19-APR-1999;	99US-0130077P.
PR 21-APR-1999;	99US-0130444P.
PR 23-APR-1999;	99US-0130510P.
PR 23-APR-1999;	99US-0130891P.
PR 28-APR-1999;	99US-0131444P.
PR 30-APR-1999;	99US-0132048P.
PR 30-APR-1999;	99US-0132407P.
PR 04-MAY-1999;	99US-0132484P.
PR 05-MAY-1999;	99US-0132485P.
PR 06-MAY-1999;	99US-0132486P.
PR 06-MAY-1999;	99US-0132487P.
PR 07-MAY-1999;	99US-0132863P.
PR 11-MAY-1999;	99US-0134256P.
PR 14-MAY-1999;	99US-0134218P.
PR 14-MAY-1999;	99US-0134219P.
PR 14-MAY-1999;	99US-0134221P.
PR 14-MAY-1999;	99US-0134370P.
PR 18-MAY-1999;	99US-0134768P.
PR 19-MAY-1999;	99US-0134941P.
PR 20-MAY-1999;	99US-0135124P.
PR 21-MAY-1999;	99US-0135353P.
PR 24-MAY-1999;	99US-0135629P.
PR 25-MAY-1999;	99US-0136021P.
PR 27-MAY-1999;	99US-0136392P.
PR 28-MAY-1999;	99US-0136782P.
PR 01-JUN-1999;	99US-0137222P.
PR 03-JUN-1999;	99US-0137528P.
PR 04-JUN-1999;	99US-0137502P.
PR 07-JUN-1999;	99US-0137724P.
PR 08-JUN-1999;	99US-0138094P.
PR 10-JUN-1999;	99US-0138540P.
PR 10-JUN-1999;	99US-0138847P.
PR 14-JUN-1999;	99US-0138119P.
PR 16-JUN-1999;	99US-0139452P.
PR 16-JUN-1999;	99US-0139453P.
PR 17-JUN-1999;	99US-0139492P.
PR 18-JUN-1999;	99US-0139454P.
PR 18-JUN-1999;	99US-0139455P.
PR 18-JUN-1999;	99US-0139456P.
PR 18-JUN-1999;	99US-0139457P.
PR 18-JUN-1999;	99US-0139458P.
PR 18-JUN-1999;	99US-0139459P.
PR 18-JUN-1999;	99US-0139460P.
PR 18-JUN-1999;	99US-0139461P.
PR 18-JUN-1999;	99US-0139462P.
PR 18-JUN-1999;	99US-0139463P.
PR 18-JUN-1999;	99US-0139750P.
PR 18-JUN-1999;	99US-0139763P.
PR 21-JUN-1999;	99US-0139817P.
PR 22-JUN-1999;	99US-0139893P.
PR 23-JUN-1999;	99US-0140353P.
PR 23-JUN-1999;	99US-0140354P.
PR 24-JUN-1999;	99US-0140695P.
PR 28-JUN-1999;	99US-0140823P.
PR 29-JUN-1999;	99US-0140991P.
PR 30-JUN-1999;	99US-0141287P.
PR 01-JUL-1999;	99US-0141842P.
PR 01-JUL-1999;	99US-0142154P.
PR 02-JUL-1999;	99US-0142055P.
PR 06-JUL-1999;	99US-0142390P.
PR 08-JUL-1999;	99US-0142803P.
PR 09-JUL-1999;	99US-0142920P.
PR 12-JUL-1999;	99US-0142977P.
PR 13-JUL-1999;	99US-0143542P.
PR 14-JUL-1999;	99US-0143624P.
PR 15-JUL-1999;	99US-0144005P.
PR 16-JUL-1999;	99US-0144085P.
PR 16-JUL-1999;	99US-0144086P.
PR 19-JUL-1999;	99US-0144325P.
PR 19-JUL-1999;	99US-0144331P.
PR 19-JUL-1999;	99US-0144332P.
PR 19-JUL-1999;	99US-0144333P.

PR 19-JUL-1999;	99US-0144334P.
PR 19-JUL-1999;	99US-0144335P.
PR 20-JUL-1999;	99US-0144352P.
PR 20-JUL-1999;	99US-0144632P.
PR 20-JUL-1999;	99US-0144884P.
PR 21-JUL-1999;	99US-0144814P.
PR 21-JUL-1999;	99US-0145086P.
PR 21-JUL-1999;	99US-0145088P.
PR 22-JUL-1999;	99US-0145085P.
PR 22-JUL-1999;	99US-0145087P.
PR 22-JUL-1999;	99US-0145089P.
PR 22-JUL-1999;	99US-0145192P.
PR 23-JUL-1999;	99US-0145145P.
PR 23-JUL-1999;	99US-0145218P.
PR 23-JUL-1999;	99US-0145224P.
PR 26-JUL-1999;	99US-0145276P.
PR 27-JUL-1999;	99US-0145913P.
PR 27-JUL-1999;	99US-0145918P.
PR 27-JUL-1999;	99US-0145919P.
PR 28-JUL-1999;	99US-0145951P.
PR 02-AUG-1999;	99US-0146386P.
PR 02-AUG-1999;	99US-0146388P.
PR 02-AUG-1999;	99US-0146389P.
PR 03-AUG-1999;	99US-0147038P.
PR 04-AUG-1999;	99US-0147204P.
PR 04-AUG-1999;	99US-0147302P.
PR 05-AUG-1999;	99US-0147192P.
PR 05-AUG-1999;	99US-0147260P.
PR 06-AUG-1999;	99US-0147303P.
PR 06-AUG-1999;	99US-0147416P.
PR 09-AUG-1999;	99US-0147493P.
PR 09-AUG-1999;	99US-0147935P.
PR 10-AUG-1999;	99US-0148171P.
PR 11-AUG-1999;	99US-0148319P.
PR 12-AUG-1999;	99US-0148341P.
PR 13-AUG-1999;	99US-0148565P.
PR 13-AUG-1999;	99US-0148684P.
PR 16-AUG-1999;	99US-0149358P.
PR 17-AUG-1999;	99US-0149175P.
PR 18-AUG-1999;	99US-0149426P.
PR 20-AUG-1999;	99US-0149722P.
PR 20-AUG-1999;	99US-0149723P.
PR 20-AUG-1999;	99US-0149929P.
PR 23-AUG-1999;	99US-0149902P.
PR 23-AUG-1999;	99US-0149930P.
PR 25-AUG-1999;	99US-0150566P.
PR 26-AUG-1999;	99US-0150884P.
PR 27-AUG-1999;	99US-0151065P.
PR 27-AUG-1999;	99US-0151066P.
PR 27-AUG-1999;	99US-0151080P.
PR 30-AUG-1999;	99US-0151303P.
PR 31-AUG-1999;	99US-0151438P.
PR 01-SEP-1999;	99US-0151930P.
PR 07-SEP-1999;	99US-0152363P.
PR 10-SEP-1999;	99US-0153070P.
PR 13-SEP-1999;	99US-0153758P.
PR 15-SEP-1999;	99US-0154018P.
PR 16-SEP-1999;	99US-0154039P.
PR 20-SEP-1999;	99US-0154779P.
PR 22-SEP-1999;	99US-0155139P.
PR 23-SEP-1999;	99US-0155486P.
PR 24-SEP-1999;	99US-0155659P.
PR 28-SEP-1999;	99US-0156458P.
PR 29-SEP-1999;	99US-0156596P.
PR 04-OCT-1999;	99US-0157117P.
PR 05-OCT-1999;	99US-0157753P.
PR 06-OCT-1999;	99US-0157865P.
PR 07-OCT-1999;	99US-0158029P.
PR 08-OCT-1999;	99US-0158232P.
PR 12-OCT-1999;	99US-0158369P.
PR 13-OCT-1999;	99US-0159293P.
PR 13-OCT-1999;	99US-0159294P.
PR 13-OCT-1999;	99US-0159295P.

PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159363P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 50.0%; Score 45; DB 3; Length 751;
Best Local Similarity 50.0%; Pred. No. 2, 2e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
5 FCIGRLCVODGF 16
84 YCIGRWCVPEPF 95

RESULT 15
AAG39903
ID AAG39903 standard; protein; 781 AA.
AC AAG39903;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 49439.
DE Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125768P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132486P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141442P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148315P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158923P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.

PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161044P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query March 50.0%; Score 45; DB 3; Length 781;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 5 FCIGRLCVODGF 16
:|:|:|:|:|
DB 114 YCLGRMCVPEPF 125

Search completed: January 23, 2006, 10:28:32
Job time : 55.1004 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:32:56 ; Search time 8.73362 seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-15
Perfect score: 90
Sequence: 1 ASFGFCIGRLCVQDGF 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	399	2 B82197	zona occludens tox
2	90	100.0	399	2 A43864	zona occludens t
3	51	56.7	245	2 T34189	hypothetical prote
4	51	56.7	371	2 H88208	protein K02A2.1 [1
5	49	54.4	440	2 E88939	protein C0584.6 [1
6	48	53.3	161	2 S61389	small basic protei
7	48	53.3	243	2 T34190	hypothetical prote
8	48	53.3	321	2 T24572	hypothetical prote
9	47	52.2	796	2 T23238	hypothetical prote
10	46.5	51.7	306	2 S47330	penicillinase - Ba
11	46	51.1	261	2 T32359	hypothetical prote
12	45	51.1	265	2 T32400	hypothetical prote
13	45	50.0	241	2 B83447	hypothetical prote
14	43	47.8	127	2 G84720	hypothetical prote
15	43	47.8	216	2 T34191	hypothetical prote
16	43	47.8	432	2 G64133	adenylosuccinate s
17	43	47.8	484	2 NNBY2	anthranilate synth
18	43	47.8	493	2 E71008	hypothetical prote
19	43	47.8	496	2 H75122	sarcosine oxidase
20	43	47.8	768	2 S04518	anthranilate synth
21	43	47.8	768	2 S07305	trpC protein - Eme
22	43	47.8	2210	1 RXXPTV	genome polypeptid
23	42.5	47.2	306	2 G69674	beta-lactamase (BC
24	42	46.7	330	2 D87068	hypothetical prote
25	42	46.7	676	2 A45984	sperm-binding glyc
26	42	46.7	745	2 A48833	kinase-binding glyc
27	42	46.7	805	2 S64238	probable anthranil
28	41	45.6	64	2 T03649	hypothetical prote
29	41	45.6	155	2 G72580	hypothetical prote

30	41	45.6	164	2 A41712	H+-exporting ATPas
31	41	45.6	256	2 T16805	hypothetical prote
32	41	45.6	276	2 UQ2340	anthranilate synth
33	41	45.6	289	2 F88939	protein C0584.5 [1
34	41	45.6	469	2 B86381	probable anthranil
35	41	45.6	762	1 NNMC2	anthranilate synth
36	40	44.4	92	2 C42507	F7L protein - Vacc
37	40	44.4	128	2 D84605	hypothetical prote
38	40	44.4	248	2 A10678	probable oxidoredu
39	40	44.4	296	2 T26723	hypothetical prote
40	40	44.4	363	2 A83177	probable N-acetyl
41	40	44.4	451	2 T24018	hypothetical prote
42	40	44.4	475	2 T01352	hypothetical prote
43	40	44.4	542	2 S49888	probable membrane
44	40	44.4	631	2 T05868	hypothetical prote
45	40	44.4	715	2 S70397	zona pellucida gly

ALIGNMENTS

RESULT 1

B82197
zona occludens toxin VCI458 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C/Spectes: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #ext_change 09-Jul-2004

C/Accession: B82197

R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.J.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragol, I.; Sellers, J.

31-R.R.T.; Makianos, J.J.; Venter, J.C.; Fraser, C.M.

Figure 406: 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen-Vibrio cholerae.

A/Reference number: A82035; WUID:20406833; PMID:10952301

A/Accession: B82197

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1339 <HEI>

A/Cross-references: UNIPROT:P38442; UNIPARC:UP1000013C456; GB:AE004224; GB:AE003852; NI

A/Experimental source: serogroup O1, strain N16961; Biotype El Tor

C/Genetics:

A/Genes: VCI458

A/Map position: 1

Query Match 100.0%; Score 90; DB 2; Length 399;

Best local Similarity 100.0%; Pred. No. 5.6e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASFGFCIGRLCVQDGF 16

DB 284 ASFGFCIGRLCVQDGF 299

RESULT 2

A43864
zona occludens toxin - Vibrio cholerae

C/Spectes: Vibrio cholerae

C/Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #ext_change 09-Jul-2004

C/Accession: A43864

R/Baudry, B.; Fasanio, A.; Kellay, J.; Kaper, J.B.

Infect. Immun. 60, 428-434, 1992

A/Title: Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.

A/Reference number: A43864; WUID:92112300; PMID:1730472

A/Accession: A43864

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1399 <BAU>

A/Cross-references: UNIPROT:P38442; UNIPARC:UP100001705A7; GB:M63563; NID:G155314; PIDN

A/Note: Sequence extracted from NCBI database (NCBIN:77488, NCBI:77491)

Query Match 100.0%; Score 90; DB 2; Length 399;

Best local Similarity 100.0%; Pred. No. 5.6e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASFGCTGRLCVQDF 16
|||||
Db 284 ASFGCTGRLCVQDF 299

RESULT 3
T34189
hypothetical protein D1022.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34189
R:Bentley, D.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid D1022.
A:Reference number: Z21487
A:Accession: T34189
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-245 <BEN>
A:Cross-references: UNIPROT:O18925; UNIPARC:UPI0000080B87; EMBL:U23517; PIDN:AAB93336.1;
A:Experimental source: strain Bristol N2; clone D1022
C:Genetics:
A:Gene: CESP:D1022.3
A:Map position: 2
A:introns: 25/2; 213/1
C:Superfamily: carboxymethylenebutenolidase

Query Match 56.7%; Score 51; DB 2; Length 245;
Best Local Similarity 69.2%; Pred. No. 0.74;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SRFGCTGRLCVQD 14
:|||||
Db 120 AFGCTGGCSTD 132

RESULT 4
H88208
protein K02A2.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: H88208
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: H88208
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <STO>
A:Cross-references: UNIPROT:Q7M3K8; UNIPARC:UPI000017A59A; GB:chr_II; PID:G726385; GSPDE
A:Note: similar to C. elegans protein B04P6.1
C:Genetics:
A:Gene: K02A2.1
A:Map position: 2
C:Superfamily: carboxymethylenebutenolidase

Query Match 56.7%; Score 51; DB 2; Length 371;
Best Local Similarity 69.2%; Pred. No. 1.1;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SRFGCTGRLCVQD 14
:|||||
Db 246 AFGCTGGCSTD 258

RESULT 5
B88939
protein C05E4.6 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: B88939
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B88939
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-440 <STO>
A:Cross-references: UNIPROT:O17358; UNIPARC:UPI000017A644; GB:chr_V; PIDN:AAB71275.1; PI
C:Genetics:
A:Gene: C05E4.6
A:Map position: 5

Query Match 54.4%; Score 49; DB 2; Length 440;
Best Local Similarity 69.2%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ASFGCTGRLCVQ 13
|||||
Db 94 ASFGCTGRLCVQ 106

RESULT 6
S61389
small basic protein sbpA - Legionella pneumophila
C:Species: Legionella pneumophila
C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S61389
R:Berger, K.H.; Merliam, J.J.; Isberg, R.R.
Mol. Microbiol. 14, 809-822, 1994
A:Title: Altered intracellular targeting properties associated with mutations in the Leg
A:Reference number: S61388; MUID:95198553; PMID:7891566
A:Accession: S61389
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <BER>
A:Cross-references: UNIPROT:Q48805; UNIPARC:UPI00000BAR24; EMBL:U07940; NID:G563802; PID
C:Genetics:
A:Gene: sbpA
C:Superfamily: carboxymethylenebutenolidase

Query Match 53.3%; Score 48; DB 2; Length 161;
Best Local Similarity 57.1%; Pred. No. 1.6;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ASFGCTGRLCVQD 14
:|||||
Db 115 AATGCTGRLCVQD 128

RESULT 7
T34190
hypothetical protein D1022.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34190
R:Bentley, D.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid D1022.
A:Reference number: Z21487
A:Accession: T34190
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-243 <BEN>
A:Cross-references: UNIPROT:O18926; UNIPARC:UPI000007C3BC; EMBL:U23517; PIDN:AAB93337.1;
A:Experimental source: strain Bristol N2; clone D1022
C:Genetics:
A:Gene: CESP:D1022.4
A:Map position: 2
A:introns: 26/2; 58/3

C/Superfamily: carboxymethylglutaminase

Query Match 53.3%; Score 48; DB 2; Length 243;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ASFGICIGRLCVQD 14
| | | | |
| | | | |
DB 119 AAIQYCFGLCTVD 132

RESULT 8

T24572
hypothetical protein T06C12.11 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T24572

R/Kelly, P.
submitted to the EMBL Data Library, October 1996

A/Reference number: Z19908

A/Accession: T24572

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-321 <WIL>

A/Cross-references: UNIPROT:O18046; UNIPARC:UPI000007B3D9; EMBL:Z81116; PIDN:CAE03307.1;

A/Experimental source: clone T06C12

C/Genetics:

A/Map position: 5 116/2; 211/1; 244/3

A/Introns: 90/1; 116/2; 211/1; 244/3

Query Match 53.3%; Score 48; DB 2; Length 321;
Best Local Similarity 59.2%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ASFGICIGRLCVQ 13
| | | | |
| | | | |
DB 94 ASFGICISILCTQ 106

RESULT 9
T23238
hypothetical protein T05H10.7 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T23238; T24560

R/Lighting, J.
submitted to the EMBL Data Library, January 1995

A/Reference number: Z19713

A/Accession: T23238

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-796 <WIL>

A/Cross-references: UNIPROT:Q10003; UNIPARC:UPI000013C01A; EMBL:Z47811; PIDN:CAA87788.1;

A/Experimental source: clone K02C4

R/Thomas, K.
submitted to the EMBL Data Library, January 1995

A/Reference number: Z19907

A/Accession: T24560

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-796 <WIL>

A/Cross-references: UNIPARC:UPI000013C01A; EMBL:Z47812; PIDN:CAA87796.1; GSPDB:GN00020;

QY 3 FGFCIGRLCVQDGF 16
| | | | |
| | | | |
DB 193 FGTVAGKRCVSDGW 206

RESULT 10

S47330
penicillinase - Bacillus amyloliquefaciens

C/Species: Bacillus amyloliquefaciens

C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Oct-2004

C/Accession: S47330

R/van Dijk, J.M.; de Jong, A.; Nauta, A.; Venema, G.; Bron, S.
submitted to the EMBL Data Library, August 1994

A/Description: Identification of penicillinase-encoding genes of Bacillus amyloliquefaciens

A/Reference number: S47330

A/Accession: S47330

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-306 <VAN>

A/Cross-references: UNIPROT:Q44674; UNIPARC:UPI00001269C9; EMBL:Z35653; NID:G522111; PII

C/Superfamily: Beta-lactamase I

Query Match 51.7%; Score 46.5; DB 2; Length 306;
Best Local Similarity 66.7%; Pred. No. 4.8;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 3 FGFCIGRLCVQ-DGF 16
| | | | |
| | | | |
DB 11 FGICIGRLCVSFTGF 25

RESULT 11

T23239
hypothetical protein T23F4.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T23239

R/Maggi, L.
submitted to the EMBL Data Library, September 1997

A/Description: The sequence of C. elegans cosmid T23F4.

A/Reference number: Z21160

A/Accession: T23239

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-261 <MAG>

A/Cross-references: UNIPROT:O17262; UNIPARC:UPI00000792AE; EMBL:AF025466; PIDN:AA871033

A/Experimental source: strain Bristol N2; clone T23F4

C/Genetics:

A/Map position: 2

A/Introns: 47/2; 82/1; 110/2; 176/2

C/Superfamily: carboxymethylglutaminase

Query Match 51.1%; Score 46; DB 2; Length 261;
Best Local Similarity 61.5%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 SFSGICIGRLCVQD 14
| | | | |
| | | | |
DB 138 SIFGFCGGLCTLD 150

RESULT 12

T23240
hypothetical protein T23F4.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T23240

R/Maggi, L.
submitted to the EMBL Data Library, September 1997

A/Description: The sequence of C. elegans cosmid T23F4.

A/Reference number: Z21160

A/Accession: T23240

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-265 <MAG>
A:Cross-references: UNIPROT:O17263; UNIPARC:UPI000007CDD7; EMBL:AF025466; PIDN:AAB71032.
A:Experimental source: strain Bristol N2; clone T23F4
C:Genetics:
A:Gene: CESP:T23F4.3
A:Map position: 2
A:Introns: 47/2; 80/3; 114/2; 180/2
C:Superfamily: carboxymethylenebutenolidase

Query Match 51.1%; Score 46; DB 2; Length 265;
Best Local Similarity 61.5%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SFPGICIGRLCYOD 14
Db 142 SIGFCFGGLCTLD 154

RESULT 13

B83447
hypothetical protein PA1597 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B83447
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Lim,
..; Lorry, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10964043
A:Accession: B83447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-241 <STO>
A:Cross-references: UNIPROT:Q913C4; UNIPARC:UPI00000C53FC; GB:AE004587; GB:AE004091; NID
C:Genetics:
A:Gene: PA1597
A:Experimental source: strain PA01
C:Superfamily: carboxymethylenebutenolidase

Query Match 50.0%; Score 45; DB 2; Length 241;
Best Local Similarity 47.4%; Pred. No. 6.9;
Matches 9; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

Qy 1 ASFGFCIGRLC---VQDG 15
Db 119 AAFGFCFGGCGCALRLARDG 137

RESULT 14

G84720
hypothetical protein At2g31440 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84720
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Motilal, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.B.; Umeyam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84720
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <STO>
A:Cross-references: UNIPARC:UPI000017A026; GB:AE002093; NID:g4589957; PIDN:AAD26475.1; G
C:Genetics:
A:Gene: At2g31440
A:Map position: 2

Query Match 47.8%; Score 43; DB 2; Length 127;

Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FCIGRLC 11
Db 51 FCIGRLC 57

RESULT 15

T34191
hypothetical protein D1022.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34191
R:Bentley, D.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid D1022.
A:Reference number: Z21487
A:Accession: T34191
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-216 <BEN>
A:Cross-references: UNIPROT:Q18927; UNIPARC:UPI00000829D0; EMBL:U23517; PIDN:AAB93338.1;
A:Experimental source: strain Bristol N2; clone D1022
C:Genetics:
A:Gene: CESP:D1022.5
A:Map position: 2
A:Introns: 26/2; 93/1; 185/1
C:Superfamily: carboxymethylenebutenolidase

Query Match 47.8%; Score 43; DB 2; Length 216;
Best Local Similarity 53.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SFPGICIGRLCYOD 14
Db 93 TLGFCFGGLCALD 105

Search completed: January 23, 2006, 10:43:08
Job time : 9.73362 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:31:19 ; Search time 51.0044 Seconds
(without alignments)
221.323 Million cell updates/sec

Title: US-10-501-838a-15
Perfect score: 90
Sequence: 1 ASFGFCIGRLCVQDGF 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_05.80:*
2: uniprot_sprot:*
3: uniprot_trembl:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	323	2	Q918F5_VIBMT
2	90	100.0	355	2	Q94M03_VIVIRU
3	90	100.0	399	1	ZOT_VIBCH
4	90	100.0	399	2	Q93A09_VIVIRU
5	90	100.0	399	2	Q94M02_VIVIRU
6	51	56.7	158	1	YRDI_CABEL
7	51	56.7	239	2	Q607F7_METCA
8	51	56.7	245	2	Q18925_CABEL
9	51	56.7	371	2	Q7M3K8_CABEL
10	49	54.4	328	2	Q17358_CABEL
11	48	53.3	161	2	Q48805_LEGPN
12	48	53.3	227	2	Q83C06_COXEU
13	48	53.3	238	2	Q5X1K4_LEGPA
14	48	53.3	238	2	Q5WTB3_LEGFL
15	48	53.3	243	2	Q18926_CABEL
16	48	53.3	248	2	Q5Z634_LEGPH
17	48	53.3	274	2	Q61C64_CABER
18	48	53.3	321	2	Q18046_CABEL
19	48	53.3	654	2	Q4PFI8_USIWA
20	47	52.2	243	2	Q61C65_CABER
21	47	52.2	427	1	LGT_MYCPA
22	47	52.2	719	2	Q7JMK7_CABEL
23	47	52.2	734	2	Q57WH2_CTRYP
24	47	52.2	796	1	YRS7_CABEL
25	47	52.2	805	2	Q61B17_CABER
26	46.5	51.7	306	1	BLAC_BACAM
27	46	51.1	261	2	Q17262_CABEL
28	46	51.1	265	2	Q17263_CABEL
29	46	51.1	265	2	Q61OAI_CABER
30	46	51.1	336	2	Q94MR7_SCAUD
31	46	51.1	931	2	Q6LK90_PHOPR

ALIGNMENTS

32	45	50.0	227	2	Q7VH89_HELHP
33	45	50.0	241	2	Q913C4_PSRAR
34	44	48.9	238	2	Q5ZS54_LEGPH
35	44	48.9	238	2	Q5X1M4_LEGPA
36	44	48.9	238	2	Q5WTD3_LEGFL
37	44	48.9	241	2	Q88FC0_PSRPK
38	44	48.9	242	2	Q4ZUM1_PSRPS
39	44	48.9	242	2	Q4KEK7_PSRPS
40	44	48.9	242	2	Q883Y8_PSRSM
41	44	48.9	400	2	Q88162_PSRPK
42	44	48.9	404	2	Q83VQ7_PPROT
43	44	48.9	441	2	Q6ZET0_SYNY3
44	44	48.9	456	2	Q4RGN6_TETNG
45	44	48.9	492	2	Q5JFD7_PYRKO
					Q7VH89 helicobacte
					Q913C4 pseudomonas
					Q5ZS54 legionella
					Q5X1M4 legionella
					Q5WTD3 legionella
					Q88FC0 pseudomonas
					Q4ZUM1 pseudomonas
					Q4KEK7 pseudomonas
					Q883Y8 pseudomonas
					Q88162 pseudomonas
					Q83VQ7 methylolact
					Q6ZET0 synechocyst
					Q4RGN6 tetradon n
					Q5JFD7 pyrococcus

RESULT 1
Q918F5_VIBMT ID Q918F5_VIBMT PRELIMINARY; PRT: 323 AA.
AC Q918F5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zot (Fragment).
GN Name=zot;
OS Vibrio mimicus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_Taxid=674;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PT5;
RX MEDLINE=20143766; PubMed=10678967;
DOI=10.1126/JAI.68.3.1507-1513.2000;
RA Boyd B.F., Moyer K.B., Shi L., Waldor M.K.;
RT "Infectious CtxPhi and the Vibrio pathogenicity island prophage in
RT Vibrio mimicus: evidence for recent horizontal transfer between V.
RT mimicus and V. cholerae."
RL Infect. Immun. 68:1507-1513(2000).
DR EMBL: AF207857; AAP40142.1; -; Genomic_DNA.
DR InterPro: IPR008900; Zot.
DR Pfam: PF05707; Zot; 1.
FT NON TER 1
FT NON TER 323
SQ SEQUENCE 323-AA: 36306 MW; 01C12DAE9B873C3B CRC64;
Query Match 100.0%; Score 90; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASFGFCIGRLCVQDGF 16
DB 233 ASFGFCIGRLCVQDGF 248
RESULT 2
Q94M03_VIVIRU ID Q94M03_VIVIRU PRELIMINARY; PRT: 355 AA.
AC Q94M03;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zot (Fragment).
GN Name=zot;
OS Vibrio phage CTX.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_Taxid=141904;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Bhattacharyya T., Nair G.B., Nandy R.K.;

```
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF414368; AAL09680.1; -; Genomic_DNA.
DR InterPro; IPR008900; Zot.
DR Pfam; PF05707; Zot; 1.
FT NON TER 355 355
SQ SEQUENCE 355 AA; 39849 MW; B77E8620AD3DB65 CRC64;

Query Match 100.0%; Score 90; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ASFGFCIGRLCVQDGF 16
|||
284 ASFGFCIGRLCVQDGF 299

RESULT 3
ZOT_VIBCH STANDARD; PRT; 399 AA.
ID ZOT_VIBCH 091706; Q9R3V6;
AC P38442; 091706; Q9R3V6;
DT 01-OCT-1994 (Rel. 30, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Zonula occludens toxin (Zonular occludens toxin) (zot).
GN Name=zot; OrderedLocNames=VCI458;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Classical Inaba 569B;
RX MEDLINE=92112300; PubMed=1730472;
RA Baudry B., Fasano A., Kelsey J., Kaper J.B.;
RT "Cloning of a gene (zot) encoding a new toxin produced by Vibrio
RT cholerae."
RL Infect. Immun. 60:428-434(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KNH002;
RX Shin H.J., Park Y.C., Kim Y.C.;
RT "Cloning and nucleotide sequence analysis of the virulence gene
RT cassette from Vibrio cholerae KNH002 isolated in Korea."
RL Misumunhag Hoiji 35:205-210(1999).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O139-Tor Ogawa;
RX Zhi-Yong H., Wei-Jie Z., Xiang-Fu W.;
RT "Cloning and Expression of zot Gene from Vibrio cholerae."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=El Tor 86015 / Serotype O1;
RX Kan B., Liu Y.Q., Qi G.M., Gao S.Y.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406633; PubMed=10952301; DOI=10.1038/35020000;
RA Heddelberg J.F., Elsen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.R., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.B., Read T.D., Tetelin H., Richardson D.L.,
RA Ernolova M.D., Vamathevan J.J., Bae S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utecher T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
RN [6]
RP CHARACTERIZATION.
RX MEDLINE=91271365; PubMed=2052603;
RA Fasano A., Baudry B., Pumplin D.W., Wasserman S.S., Tall B.D.,
```

```
RA Kelsey J.M., Kaper J.B.;
RT "Vibrio cholerae produces a second enterotoxin, which affects
RT intestinal tight junctions."
RL Proc. Natl. Acad. Sci. U.S.A. 88:5242-5246(1991).
CC -1- FUNCTION: Increases the permeability of the small intestine
CC by affecting the structure of intercellular tight junctions
CC (zonula occludens).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; M83563; AAA27582.1; -; Genomic_DNA.
DR EMBL; AF175708; AAD51358.1; -; Genomic_DNA.
DR EMBL; AF123049; AAD26854.1; -; Genomic_DNA.
DR EMBL; AF220606; AAF29547.1; -; Genomic_DNA.
DR EMBL; AE004224; AAF94615.1; -; Genomic_DNA.
DR PIR; A43864; A43864.
DR PIR; B82197; B82197.
DR TIGR; VCI458; -.
DR InterPro; IPR008900; Zot.
DR Pfam; PF05707; Zot; 1.
KW Complete proteome; Enterotoxin; Toxin.
FT VARIANT 45 45 M -> I (in strain 569B).
FT VARIANT 100 100 V -> A (in strain 569B).
FT VARIANT 272 272 V -> A (in strain 569B).
FT VARIANT 281 281 V -> A (in strain 569B).
FT VARIANT 349 349 A -> S (in strain 86015).
FT VARIANT 381 381 K -> R (in strain 86015).
FT CONFLICT 386 399 IKTENDKKGLNSIF -> VKKXKESIKSPL (in Ref.
SQ SEQUENCE 399 AA; 44903 MW; 3C7424B758176774 CRC64;

Query Match 100.0%; Score 90; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ASFGFCIGRLCVQDGF 16
|||
284 ASFGFCIGRLCVQDGF 299

RESULT 4
ID Q93A09_VIVRU PRELIMINARY; PRT; 399 AA.
AC Q93A09;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Zot.
GN Name=zot;
OS Vibrio phage CTX.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=141904;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Rul Y., Kan B., Gao S., Liu Y., Qi G.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF416590; AAL09690.1; -; Genomic_DNA.
DR InterPro; IPR008900; Zot.
DR Pfam; PF05707; Zot; 1.
SQ SEQUENCE 399 AA; 45007 MW; 057698567160539B CRC64;

Query Match 100.0%; Score 90; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ASFGFCIGRLCVQDGF 16
|||
284 ASFGFCIGRLCVQDGF 299
```



```

RESULT 5
094M02_GVIRU PRELIMINARY; PRT; 399 AA.
AC 094M02;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
DE Zot;
GN Name=Zot;
OS Vibrio phage CTX.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=141904;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Bhattacharyya T., Nandy R.K., Nair G.B.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Bhattacharyya T., Nandy R.K., Nair G.B.;
RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF413369; AAL09684.1; -; Genomic_DNA.
DR EMBL; AF542088; AAN17344.1; -; Genomic_DNA.
DR EMBL; AF542089; AAN17350.1; -; Genomic_DNA.
DR InterPro; IPR008900; Zot.
DR Pfam; PF05707; Zot; 1.
SQ SEQUENCE 399 AA; 44902 MW; 367EF3430F00CB73 CRC64;

Query Match 100.0%; Score 90; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 6; 6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASFGFCIGRLCYQDGF 16
Db 284 ASFGFCIGRLCYQDGF 299

RESULT 6
YR01_CABEL STANDARD; PRT; 158 AA.
AC 009571;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical protein K02A2.1 in chromosome II.
GN OlfAmees-K02A2.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA STRAIN=Br1etel N2;
RX MEDLINE=99069613; Pubmed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE REVISION.
RG WormBase consortium;
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U23171; AAC46705.2; -; Genomic_DNA.
CC DR Ensemble; K02A2.1; Caenorhabditis elegans.
CC WormBase; WBGene00019288; K02A2.1.

```

```

DR WormPep; K02A2.1; CR21012.
DR InterPro; IPR002925; DieneLactn_hydro.
DR Pfam; PF01738; DLH; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 158 AA; 17251 MW; DD0456ADC388492 CRC64;

Query Match 56.7%; Score 51; DB 1; Length 158;
Best Local Similarity 69.2%; Pred. No. 4; 6;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SFPGFCIGRLCYQD 14
Db 33 ARFGFCIGRLCYQD 45

RESULT 7
0607F7_METCA PRELIMINARY; PRT; 239 AA.
AC 0607F7;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE DieneLactone hydrolase family protein.
GN OrderedCusNames-MCA1803;
OS Methylococcus capsulatus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
OC Methylococcaceae; Methylococcus.
OX NCBI_TaxID=414;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=Bath / NCIMB 11132;
RX Pubmed=15383840; DOI=10.1371/journal.pbio.0020303;
RA Ward N.L., Larsen O., Sakwa J., Bruseck L., Knouri H.M., Durkin A.S.,
RA Dimitrov G., Jiang L., Scanlan D., Kang K.H., Lewis M.R., Nelson K.E.,
RA Mehe B.A., Wu M., Heidelberg J.F., Paulsen I.T., Fouts D.E.,
RA Ravel J., Tettelin H., Ren O., Read T.D., DeBoy R.T., Seshadri R.,
RA Salzberg S.L., Jensen H.B., Birkeland N.K., Nelson W.C., Dodson R.J.,
RA Grindhaug S.H., Holt I.E., Sidhammer I., Jonassen I., Vanaken S.,
RA Uterback T.R., Feldblyum T.V., Fraser C.M., Lilienhaug J.R.,
RA Bisen J.A.;
RT "Genomic insights into methanotrophy: the complete genome sequence of
RT Methylococcus capsulatus (Bath).";
RL Plos Biol. 2:1616-1628(2004).
CC -1- FUNCTION: Ring cleavage of cyclic ester dieneLactone to produce
CC maleylacetate (By similarity).
CC -1- CATALYTIC ACTIVITY: 4-carboxymethylmethyl-2-en-4-olide + H(2)O =
CC 4-oxohex-2-enedioate.
CC EMBL; AB017282; AAU92176.1; -; Genomic_DNA.
DR TIGR; MCA1803; -;
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002925; DieneLactn_hydro.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF01738; DLH; 1.
KW Aromatic hydrocarbons catabolism; Complete proteome; Hydrolase;
KW Serine esterase.
SQ SEQUENCE 239 AA; 26067 MW; 8586B8F2B737A47 CRC64;

Query Match 56.7%; Score 51; DB 2; Length 239;
Best Local Similarity 64.3%; Pred. No. 6; 7;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ASFGFCIGRLCYQD 14
Db 116 AAMGFCIGRLCYQD 129

RESULT 8
Q18925_CABEL PRELIMINARY; PRT; 245 AA.
AC 018925;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

```

AC	017358;
DT	01-JAN-1998 (TREMBLrel. 05, Created)
DT	01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Seven tm receptor protein 134.
GN	Name=atr-134; ORFNames=C05E4.6;
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdida; Rhabditoidea;
OC	Rhabditiidae; Peloderinae; Caenorhabditis.
OX	NCBI_TaxID=6239;
RN	[1]
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC	STRAIN-Bristol N2;
RX	MEDLINE=99069613; PubMed=9851916;
RG	The C. elegans sequencing consortium;
RT	"Genome sequence of the nematode C. elegans: a platform for
RL	investigating biology."
RU	Science 282:2012-2018(1998).
DR	EMBL AF026209; AAB71275.2; -; Genomic_DNA.
DR	PIR; B88939; B88939.
DR	Ensembl; C05E4.6; Caenorhabditis elegans.
DR	WormBase; WBGene0006182; C05E4.6.
DR	WormPep; C05E4.6; CR25759.
DR	GO; GO:0016020; C:membrane; IEA.
DR	GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR	GO; GO:0004872; F:receptor activity; IEA.
DR	InterPro; IPR003002; 7TM_Chemreceptl.
DR	InterPro; IPR000168; Nm7TM_Chemrecept.
KW	Pfam; PF01461; Tcm_4; 1.
SQ	Complete proteome; Receptor.
	SEQUENCE 328 AA; 37435 MW; 56F04DB138A4649 CRC64;
Query Match	54.4%; Score 49; DB 2; Length 328;
Best Local Similarity	69.2%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	1 ASFGICRLCTQ 13
Db	94 ASFGICSLCTQ 106
RESULT 11	
Q48805_LEGN	
ID Q48805_LEGN PRELIMINARY;	PRT; 161 AA.
AC Q48805;	
DT 01-NOV-1996 (TREMBLrel. 01, Created)	
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)	
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
DE SbpA.	
GN Name=sbpa;	
OS Legionella pneumophila.	
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;	
OC Legionellaceae; Legionella.	
OX NCBI_TaxID=446;	
RN [1]	
RP NUCLEOTIDE SEQUENCE.	
RC STRAIN=LPO2;	
RX MEDLINE=95198553; PubMed=7891566;	
RA Berger K.H., Meriam J.J., Isberg R.R.;	
RT "Altered intracellular targeting properties associated with mutations	
in the Legionella pneumophila dotA gene."	
RL Mol. Microbiol. 14:809-822(1994).	
-!- FUNCTION: Ring cleavage of cyclic ester diene lactone to produce	
methylacetate (by similarity).	
CC -!- CARBONIC ACTIVITY: 4-carboxymethylenebut-2-en-4-olide + H(2)O =	
CC 4-oxohex-2-enedioate.	
CC EMBL; U07940; AAA79903.1; -; Genomic_DNA.	
DR PIR; S61389; S61389.	
DR GO; GO:0016787; F:hydrolase activity; IEA.	
DR InterPro; IPR002925; DieneIactn_hydro.	
DR InterPro; IPR000379; Set_estrs.	
DR Pfam; PF01738; DUF; 1.	
KW Aromatic hydrocarbons catabolism; Hydrolase; Serine esterase.	

SQ SEQUENCE 161 AA; 17324 MW; B59D2591B1614437 CRC64;
 Query Match 53.3%; Score 48; DB 2; Length 161;
 Best Local Similarity 57.1%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ASFGFCIGRLCYQD 14
 Db 115 AAIQYCFGLCYVD 128
 RESULT 12
 Q83CJ6 COXBU PRELIMINARY; PRT; 237 AA.
 AC Q83CJ6;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DE dienelectone hydrolase family protein.
 GN OrderedLocustNames=CBU1119;
 OS Coccidia burnetii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 CC Coccidiaceae; Coccidia.
 OK NCBI_TaxId=7771;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=Nine Mile phase 1 / RSA 493;
 RC MEDLINE=2260657; PubMed=12704232; DOI=10.1073/pnas.0931379100;
 RA Sehadati R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
 RA Nelson W.C., Ward N.D., Tectelin H., Davidson T.M., Benan M.J.,
 RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
 RA Khouri H.M., Lee K.H., Cary H.A., Scanlan D., Heitzen R.A.,
 RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
 RT "Complete genome sequence of the Q-fever pathogen, Coccidia
 burnetii.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
 CC -1- FUNCTION: Ring cleavage of cyclic ester dienelectone to produce
 maleylacetate (By similarity).
 CC -1- CATALYTIC ACTIVITY: 4-carboxymethylenebut-2-en-4-olide + H(2)O =
 4-oxohex-3-enedioate.
 CC EMBL, AB016963; MAF09632.1; -; Genomic_DNA.
 DR TIGR; CBU1119;
 DR GO; GO:0016787, F:hydrolase activity; IEA.
 DR InterPro; IPR002925; Dienelectn_hydro.
 DR Pfam; PF01738; DLH; 1.
 KW Aromatic hydrocarbons catabolism; Complete proteome; Hydrolase;
 KM Serine esterase.
 SQ SEQUENCE 237 AA; 26492 MW; 365CB214F373FDB0 CRC64;
 Query Match 53.3%; Score 48; DB 2; Length 237;
 Best Local Similarity 57.1%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ASFGFCIGRLCYQD 14
 Db 114 AAIQYCFGLCYVD 127
 RESULT 13
 Q5X1K4 LEGPA PRELIMINARY; PRT; 238 AA.
 AC Q5X1K4;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DE Hypothetical protein sbpa.
 GN Name=sbpa; OrderedLocustNames=lp2739;
 OS Legionella pneumophila (strain Paris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 CC Legionellaceae; Legionella.
 OK NCBI_TaxId=297246;
 RN (1)

RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15467720; DOI=10.1038/ng1447;
 RA Cazale C., Rusnok C., Brueggemann H., Zidane N., Magnier A., Ma L.,
 RA Ticht M., Jarraud S., Bouchier C., Vandenesch F., Kunst F.,
 RA Etienne J., Glaser P., Buchrieser C.;
 RT "Evidence in the Legionella pneumophila genome for exploitation of
 host cell functions and high genome plasticity.";
 RL Nat. Genet. 36:1165-1173(2004).
 DR EMBL; CR628336; CAH1892.1; -; Genomic_DNA.
 DR Legioliast; lpp2739;
 DR GO; GO:0016787, F:hydrolase activity; IEA.
 DR InterPro; IPR002925; Dienelectn_hydro.
 DR Pfam; PF01738; DLH; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 238 AA; 26411 MW; 77C58C6B1C77D90 CRC64;
 Query Match 53.3%; Score 48; DB 2; Length 238;
 Best Local Similarity 57.1%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ASFGFCIGRLCYQD 14
 Db 115 AAIQYCFGLCYVD 128
 RESULT 14
 Q5WTB3 LEGPL PRELIMINARY; PRT; 238 AA.
 AC Q5WTB3;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DE Hypothetical protein sbpa.
 GN Name=sbpa; OrderedLocustNames=lp12612;
 OS Legionella pneumophila (strain Lens).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 CC Legionellaceae; Legionella.
 OK NCBI_TaxId=297245;
 RN NUCLEOTIDE SEQUENCE.
 RP PubMed=15467720; DOI=10.1038/ng1447;
 RA Cazale C., Rusnok C., Brueggemann H., Zidane N., Magnier A., Ma L.,
 RA Ticht M., Jarraud S., Bouchier C., Vandenesch F., Kunst F.,
 RA Etienne J., Glaser P., Buchrieser C.;
 RT "Evidence in the Legionella pneumophila genome for exploitation of
 host cell functions and high genome plasticity.";
 RL Nat. Genet. 36:1165-1173(2004).
 CC -1- FUNCTION: Ring cleavage of cyclic ester dienelectone to produce
 maleylacetate (By similarity).
 CC -1- CATALYTIC ACTIVITY: 4-carboxymethylenebut-2-en-4-olide + H(2)O =
 4-oxohex-2-enedioate.
 CC EMBL; CR628337; CAH1853.1; -; Genomic_DNA.
 DR Legioliast; lp12612;
 DR GO; GO:0016787, F:hydrolase activity; IEA.
 DR InterPro; IPR002925; Dienelectn_hydro.
 DR Pfam; PF01738; DLH; 1.
 KW Aromatic hydrocarbons catabolism; Complete proteome; Hydrolase;
 KM Hypothetical protein; Serine esterase.
 SQ SEQUENCE 238 AA; 26357 MW; 2D332F0503E1E27 CRC64;
 Query Match 53.3%; Score 48; DB 2; Length 238;
 Best Local Similarity 57.1%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ASFGFCIGRLCYQD 14
 Db 115 AAIQYCFGLCYVD 128
 RESULT 15
 Q18926 CAEBL PRELIMINARY; PRT; 243 AA.
 ID Q18926 CAEBL

AC 018926; [1]
 DT 01-NOV-1996 (TrEMBLrel_01, Created)
 DT 01-NOV-1996 (TrEMBLrel_01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel_25, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=D1022.4;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Briscoll N2;
 RX MEDLINE=9906613; PubMed=9851916; DOI=10.1126/science.282.5396.2012
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 CC -1- FUNCTION: Ring cleavage of cyclic ester dieneolone to produce
 CC maleylacetate (by similarity).
 CC -1- CATALYTIC ACTIVITY: 4-carboxymethylenebut-2-en-4-olide + H(2)O =
 CC 4-oxohe-2-enedioate.
 CC EMBL, U3517; AAB9337.1; -; Genomic DNA.
 DR PIR, T34190; T34190.
 DR Ensemble; D1022.4; Caenorhabditis elegans.
 DR WormBase; WBGenome0010723; D1022.4.
 DR Wormpep; D1022.4; CE02578.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR002925; D1eneactn_hydro.
 DR InterPro; IPR000379; Ser_eatrs.
 DR Pfam; PF01738; DLM; 1.
 KW Aromatic hydrocarbons catabolism; Complete proteome; Hydrolase;
 KW Hypothetical protein; Serine esterase.
 SQ SEQUENCE 243 AA; 26733 MW; C4AF277FBB42D0FE CRC64;

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 09:35:50 ; Search time 14.6725 seconds
(without alignments)
90.156 Million cell updates/sec

Title: US-10-501-838a-15
Perfect score: 90
Sequence: 1 ASFGFCIGRLCVQDGF 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents, AA:*
1: /cgn2_6/prodata/1/1aa/5-COMB.pep.*
2: /cgn2_6/prodata/1/1aa/6-COMB.pep.*
3: /cgn2_6/prodata/1/1aa/H-COMB.pep.*
4: /cgn2_6/prodata/1/1aa/ECTUS-COMB.pep.*
5: /cgn2_6/prodata/1/1aa/RE-COMB.pep.*
6: /cgn2_6/prodata/1/1aa/Backlist1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	400	1	US-08-624-601-8
2	70	77.8	14	2	US-09-127-815D-38
3	70	77.8	14	2	US-10-095-450-38
4	70	77.8	14	2	US-10-648-642-38
5	45	50.0	8	2	US-09-127-815D-6
6	45	50.0	8	2	US-10-095-450-6
7	45	50.0	8	2	US-10-648-642-6
8	45	50.0	240	2	US-09-252-991A-26041
9	43	47.8	250	2	US-09-270-767-36647
10	43	47.8	250	2	US-09-270-767-51864
11	43	47.8	349	2	US-09-248-796A-18207
12	42.5	47.2	84	1	US-08-465-380-51
13	42.5	47.2	84	1	US-08-486-397-51
14	42.5	47.2	84	1	US-08-486-399-51
15	42.5	47.2	84	1	US-08-461-965-51
16	42.5	47.2	84	1	US-08-634-641-51
17	42.5	47.2	84	2	US-09-249-471-51
18	42.5	47.2	84	2	US-09-249-472-51
19	42.5	47.2	84	2	US-09-249-472-51
20	42.5	47.2	84	2	US-09-249-472-51
21	42.5	47.2	84	2	US-08-809-455-51
22	42.5	47.2	84	2	US-09-249-461-51
23	42.5	47.2	84	2	US-09-249-448-51
24	42.5	47.2	84	2	US-09-249-473-51
25	42.5	47.2	162	1	US-08-498-556-51
26	42.5	47.2	162	1	US-08-465-380-64
27	42.5	47.2	162	1	US-08-486-397-64

ALIGNMENTS

28	42.5	47.2	162	1	US-08-461-965-64	Sequence 64, Appl
29	42.5	47.2	162	1	US-08-634-641-64	Sequence 64, Appl
30	42.5	47.2	162	2	US-09-249-471-64	Sequence 64, Appl
31	42.5	47.2	162	2	US-09-249-472-64	Sequence 64, Appl
32	42.5	47.2	162	2	US-09-249-451-64	Sequence 64, Appl
33	42.5	47.2	162	2	US-08-809-455-64	Sequence 64, Appl
34	42.5	47.2	162	2	US-09-249-461-64	Sequence 64, Appl
35	42.5	47.2	162	2	US-09-249-448-64	Sequence 64, Appl
36	42.5	47.2	162	2	US-09-249-473-64	Sequence 64, Appl
37	42.5	47.2	162	2	US-09-498-556-64	Sequence 64, Appl
38	42.5	47.2	8	2	US-10-127-815D-2	Sequence 2, Appl
39	42.5	47.2	8	2	US-10-095-450-2	Sequence 2, Appl
40	42.5	47.2	8	2	US-10-648-642-2	Sequence 2, Appl
41	42.5	47.2	745	1	US-08-453-947-5	Sequence 5, Appl
42	42.5	47.2	745	1	US-08-038-948-9	Sequence 5, Appl
43	42.5	47.2	745	1	US-08-453-952-5	Sequence 5, Appl
44	42.5	47.2	745	1	US-08-484-993B-43	Sequence 43, Appl
45	42.5	47.2	745	1	US-08-862-903-5	Sequence 5, Appl

RESULT 1
US-08-624-601-8

Sequence 87 Application US/08624601
Patent No. 5882653

GENERAL INFORMATION:

APPLICANT: Kaper Dr., James B.

TITLE OF INVENTION: Vibrio cholerae O1 (CVD11) and non-O1

TITLE OF INVENTION: (CVD11 and CVD12) serogroup vaccine strains, methods

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

STREET: Spencer & Frank

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/624,601

FILING DATE: 08-APR-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Schellier Dr., John W.

REGISTRATION NUMBER: 26,031

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 414-4000

TELEFAX: (202) 414-4040

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 400 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHEICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Vibrio cholerae

STRAIN: El Tor 7946

IMMEDIATE SOURCE:

CLONE: zot

US-08-624-601-8

Query Match 100.0%; Score 90; DB 1; Length 400;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASFGFCIGRLCVQDGF 16
|||
Db 284 ASFGFCIGRLCVQDGF 299

RESULT 2
US-09-127-815D-38
Sequence 38, Application US/09127815D
Patent No. 6458925

GENERAL INFORMATION:

APPLICANT: PASANO, Aleesio
TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
FILE REFERENCE: A7242

CURRENT APPLICATION NUMBER: US/09/127,815D
CURRENT FILING DATE: 1998-08-03

NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 38

LENGTH: 14
TYPE: PRT
ORGANISM: Zonulin Occludes Toxin
US-09-127-815D-38

Query Match 77.8%; Score 70; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FCIGRLCVQDGF 16
|||
Db 1 FCIGRLCVQDGF 12

RESULT 3
US-10-095-450-38

Sequence 38, Application US/10095450
Patent No. 6670448

GENERAL INFORMATION:
APPLICANT: PASANO, Aleesio
TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
FILE REFERENCE: A7242

CURRENT APPLICATION NUMBER: US/10/095,450
CURRENT FILING DATE: 2002-03-13

PRIOR APPLICATION NUMBER: 09/127,815
PRIOR FILING DATE: 1998-08-03

NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 38

LENGTH: 14
TYPE: PRT
ORGANISM: Zonulin Occludes Toxin
US-10-095-450-38

Query Match 77.8%; Score 70; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FCIGRLCVQDGF 16
|||
Db 1 FCIGRLCVQDGF 12

RESULT 4
US-10-648-642-38

Sequence 38, Application US/10648642
Patent No. 6936689

GENERAL INFORMATION:
APPLICANT: PASANO, Aleesio

TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
FILE REFERENCE: A7242

CURRENT APPLICATION NUMBER: US/10/648,642
CURRENT FILING DATE: 2003-08-27

PRIOR APPLICATION NUMBER: US/10/095,450
PRIOR FILING DATE: 2002-03-13

PRIOR APPLICATION NUMBER: 09/127,815
PRIOR FILING DATE: 1998-08-03

NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 38

LENGTH: 14
TYPE: PRT
ORGANISM: Zonulin Occludes Toxin
US-10-648-642-38

Query Match 77.8%; Score 70; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FCIGRLCVQDGF 16
|||
Db 1 FCIGRLCVQDGF 12

RESULT 5
US-09-127-815D-6

Sequence 6, Application US/09127815D
Patent No. 6458925

GENERAL INFORMATION:
APPLICANT: PASANO, Aleesio
TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
FILE REFERENCE: A7242

CURRENT APPLICATION NUMBER: US/09/127,815D
CURRENT FILING DATE: 1998-08-03

NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6

LENGTH: 8
TYPE: PRT
ORGANISM: Zonulin Peptide Antagonist
US-09-127-815D-6

Query Match 50.0%; Score 45; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GRLCVQDG 15
|||
Db 1 GRLCVQDG 8

RESULT 6
US-10-095-450-6

Sequence 6, Application US/10095450
Patent No. 6670448

GENERAL INFORMATION:
APPLICANT: PASANO, Aleesio
TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
FILE REFERENCE: A7242

CURRENT APPLICATION NUMBER: US/10/095,450
CURRENT FILING DATE: 2002-03-13

PRIOR APPLICATION NUMBER: 09/127,815
PRIOR FILING DATE: 1998-08-03

NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6

LENGTH: 8
TYPE: PRT
ORGANISM: Zonulin Peptide Antagonist

US-10-095-450-6

Query Match 50.0%; Score 45; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GRLCVQDG 15
| | | | |
| | | | |
DB 1 GRLCVQDG 8

RESULT 7
US-10-648-642-6
Sequence 6, Application US/10648642

Patent No. 6936689
GENERAL INFORMATION:
APPLICANT: PASANO, Alessandro
TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
FILE REFERENCE: A7242
CURRENT APPLICATION NUMBER: US/10/648,642
CURRENT FILING DATE: 2003-08-27
PRIOR APPLICATION NUMBER: US/10/095,450
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 09/127,815
PRIOR FILING DATE: 1998-08-03
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6:
LENGTH: 8
TYPE: PRT
ORGANISM: Zonulin Peptide Antagonist
US-10-648-642-6

Query Match 50.0%; Score 45; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GRLCVQDG 15
| | | | |
| | | | |
DB 1 GRLCVQDG 8

RESULT 8
US-09-252-991A-26041
Sequence 26041, Application US/09252991A

Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: MARC J. RUBENFELD et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26041
LENGTH: 240
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26041

Query Match 50.0%; Score 45; DB 2; Length 240;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 ASFGFCIGRLC---VQDG 15
| : | | | | : | | | | |
| : | | | | : | | | | |
DB 139 AAFGFCFGCCALBLARDG 157

RESULT 9
US-09-270-767-36647
Sequence 36647, Application US/09270767

Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 36647
LENGTH: 250
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-36647

Query Match 47.8%; Score 43; DB 2; Length 250;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 FGFCIGRLCVQDGF 16
| | | | | : | | | | |
| | | | | : | | | | |
DB 75 FGFCIGRLCVQDGF 88

RESULT 10
US-09-270-767-51864
Sequence 51864, Application US/09270767

Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 51864
LENGTH: 250
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-51864

Query Match 47.8%; Score 43; DB 2; Length 250;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 FGFCIGRLCVQDGF 16
| | | | | : | | | | |
| | | | | : | | | | |
DB 75 FGFCIGRLCVQDGF 88

RESULT 11
US-09-248-796A-18207
Sequence 18207, Application US/09248796A

Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 18207
LENGTH: 349
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: UNSURE
LOCATION: (345)
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-18207

Query Match 47.8%; Score 43; DB 2; Length 349;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 FGFCIGRLCVQDGF 16
|||:|:|:
Db 112 FGVCWGCCECFDVF 125

RESULT 12

US-08-465-380-51
Sequence 51, Application US/08465380
Patent No. 5863894

GENERAL INFORMATION:

APPLICANT: George P. Vlasuk, Patric H. Stanssens,
Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Ganssemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,380
FILING DATE: June 5, 1995
CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-465-380-51

Query Match 47.2%; Score 42.5; DB 1; Length 84;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 6 CIGRLCVQD-GF 16
|||:|:|:
Db 46 CIGRVCCDEGF 57

RESULT 13

US-08-486-397-51
Sequence 51, Application US/08486397
Patent No. 5866542

GENERAL INFORMATION:

APPLICANT: George P. Vlasuk, Patric H. Stanssens,
Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Ganssemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 357
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/269
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-51

Query Match 47.2%; Score 42.5; DB 1; Length 84;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 6 CIGRLCVQD-GF 16
|||:|:|:
Db 46 CIGRVCCDEGF 57

RESULT 14

US-08-486-399-51
Sequence 51, Application US/08486399


```

; Patent No. 5866543
; GENERAL INFORMATION:
; APPLICANT: George P. Vlausk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Ganssems, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,399
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-486-399-51

Query Match          47.2%; Score 42.5; DB 1; Length 84;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      6 CIGRLCVQD-GF 16
        |||:|||||
DB      46 CIGRVCVDEGF 57

RESULT 15
US-08-461-965-51
; Sequence 51, Application US/08461965
; Patent No. 5872098
; GENERAL INFORMATION:
; APPLICANT: George P. Vlausk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Ganssems, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon

```

```

; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,965
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 210/243
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-461-965-51

Query Match          47.2%; Score 42.5; DB 1; Length 84;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      6 CIGRLCVQD-GF 16
        |||:|||||
DB      46 CIGRVCVDEGF 57

Search completed: January 23, 2006, 10:46:49
Job time : 15.6725 secs

```

This Page Blank (1/1)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:52:43 ; Search time 36.7511 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-15
Perfect score: 90
Sequence: 1 ASFGFCIGRLCYQDGF 16

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 41782326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubppa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	100.0	16	US-10-665-184-15	Sequence 15, Appl
2	90	100.0	16	US-10-942-300-15	Sequence 15, Appl
3	90	100.0	16	US-10-501-838a-15	Sequence 15, Appl
4	90	100.0	16	US-10-891-492-1	Sequence 1, Appl
5	70	77.8	14	US-10-095-450-38	Sequence 38, Appl
6	70	77.8	14	US-10-648-642-38	Sequence 38, Appl
7	70	77.8	14	US-10-275-877-33	Sequence 33, Appl
8	70	77.8	14	US-11-074-727a-38	Sequence 38, Appl
9	70	77.8	14	US-11-074-727a-38	Sequence 38, Appl
10	48	53.3	238	US-10-282-122a-61351	Sequence 61351, A
11	47	52.2	129	US-10-424-599-166803	Sequence 166803, A
12	45	50.0	8	US-10-095-450-6	Sequence 6, Appl
13	45	50.0	8	US-10-648-642-6	Sequence 6, Appl
14	45	50.0	8	US-10-275-877-6	Sequence 6, Appl
15	45	50.0	8	US-11-074-727a-6	Sequence 6, Appl
16	45	50.0	8	US-10-424-599-274778	Sequence 274778, A
17	44	48.9	72	US-10-425-115-27448	Sequence 27448, A
18	44	48.9	111	US-10-466-242-35	Sequence 35, Appl
19	44	48.9	111	US-10-425-114-71058	Sequence 71058, A
20	44	48.9	213	US-10-425-115-288591	Sequence 288591, A
21	44	48.9	238	US-10-282-122a-61296	Sequence 61296, A
22	44	48.9	525	US-10-108-260a-4768	Sequence 4768, Ap
23	44	48.9	839	US-10-425-115-312742	Sequence 312742, A
24	43.5	48.3	552	US-10-425-115-357163	Sequence 357163, A
25	43	47.8	83	US-10-435-115-216823	Sequence 216823, A
26	43	47.8	83	US-10-156-761-12893	Sequence 12893, A
27	43	47.8	432	US-09-815-242-11264	Sequence 11264, A

28	43	47.8	432	4	US-10-282-122a-58565	Sequence 58565, A
29	43	47.8	432	4	US-10-282-122a-67137	Sequence 67137, A
30	43	47.8	520	4	US-10-032-585-7865	Sequence 7865, Ap
31	43	47.8	944	5	US-10-128-558-231	Sequence 231, Ap
32	42.5	47.2	84	3	US-09-498-272-51	Sequence 51, Appl
33	42.5	47.2	84	6	US-11-092-336-51	Sequence 51, Appl
34	42.5	47.2	101	4	US-10-437-963-102532	Sequence 102532, A
35	42.5	47.2	162	3	US-09-498-272-64	Sequence 64, Appl
36	42.5	47.2	162	6	US-11-092-336-64	Sequence 64, Appl
37	42	46.7	8	4	US-10-095-450-2	Sequence 2, Appl
38	42	46.7	8	4	US-10-648-642-2	Sequence 2, Appl
39	42	46.7	8	5	US-10-275-877-2	Sequence 2, Appl
40	42	46.7	63	5	US-10-721-793-80	Sequence 80, Appl
41	42	46.7	76	4	US-10-425-114-62453	Sequence 62453, A
42	42	46.7	85	5	US-10-721-793-78	Sequence 78, Appl
43	42	46.7	87	4	US-10-437-963-111051	Sequence 111051, A
44	42	46.7	210	5	US-10-774-355a-2378	Sequence 2378, Ap
45	42	46.7	210	5	US-10-774-355a-2378	Sequence 2378, Ap

ALIGNMENTS

RESULT 1
US-10-665-184-15
; Sequence 15, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Binat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 15
; LENGTH: 16
; TYPE: PRT
; ORGANISM: zomula occludens toxin
US-10-665-184-15

Query Match 100.0%; Score 90; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASFGFCIGRLCYQDGF 16
DB 1 ASFGFCIGRLCYQDGF 16

RESULT 2
US-10-942-300-15
; Sequence 15, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Binat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: Patentin version 3.2
;; SEQ ID NO 15
;; LENGTH: 16
;; TYPE: PRT
;; ORGANISM: zonula occludens toxin
US-10-942-300-15

Query Match 100.0%; Score 90; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASFGFCIGRLCYQDGF 16
Db 1 ASFGFCIGRLCYQDGF 16

RESULT 3

US-10-501-838A-15
; Sequence 15, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Eilat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 15
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Zonula occludens toxin
US-10-501-838A-15

Query Match 100.0%; Score 90; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASFGFCIGRLCYQDGF 16
Db 1 ASFGFCIGRLCYQDGF 16

RESULT 4

US-10-891-492-1
; Sequence 1, Application US/10891492
; Publication No. US20050059593A1
; GENERAL INFORMATION:
; APPLICANT: Fasano, Alessio
; APPLICANT: Vogel, Stefanie
; TITLE OF INVENTION: AGONIST POLYPEPTIDE OF RECEPTOR FOR ZOT
; FILE REFERENCE: 006552.00097
; CURRENT APPLICATION NUMBER: US/10/891,492
; PRIOR FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: 60/487,889
; PRIOR FILING DATE: 2003-07-15
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Vibrio Cholerae phage CTXph1
US-10-891-492-1

Query Match 100.0%; Score 90; DB 5; Length 399;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASFGFCIGRLCYQDGF 16
Db 284 ASFGFCIGRLCYQDGF 299

RESULT 5

US-10-095-450-38
; Sequence 38, Application US/10095450
; Publication No. US20020115825A1
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
; FILE REFERENCE: A7242
; CURRENT APPLICATION NUMBER: US/10/095,450
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/127,815
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Zonulin Occludes Toxin
US-10-095-450-38

Query Match 77.8%; Score 70; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FCIGRLCYQDGF 16
Db 1 FCIGRLCYQDGF 12

RESULT 6

US-10-648-642-38
; Sequence 38, Application US/10648642
; Publication No. US20040082512A1
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
; FILE REFERENCE: A7242
; CURRENT APPLICATION NUMBER: US/10/648,642
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/10/095,450
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/127,815
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Zonulin Occludes Toxin
US-10-648-642-38

Query Match 77.8%; Score 70; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 FCIGRLCYQDGF 16
Db 1 FCIGRLCYQDGF 12

RESULT 7

US-10-275-877-33

```

NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 112
TYPE: PRT
ORGANISM: Vibrio cholerae phage CTXph1
US-10-891-492-6

Query Match          77.8% Score 70; DB 5; Length 112;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5  FCIGRLCYDQGF 16
        |||||
        1 FCIGRLCYDQGF 12

RESULT 10
US-10-282-122A-61351
Sequence 61351, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: EPITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/255,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PAM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61351
LENGTH: 238
TYPE: PRT
ORGANISM: Legionella pneumophila
US-10-282-122A-61351

Query Match          53.3% Score 48; DB 4; Length 238;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1  ASFGFCIGRLCYOD 14
        |:::|::|::|

```

DB 115 AAGYCFGLCYLD 128

RESULT 11

US-10-424-599-166803
; Sequence 166803, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: LA ROSA Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5323)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166803
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121638C.1.pep
US-10-424-599-166803

Query Match 52.2%; Score 47; DB 4; Length 129;
Best Local Similarity 63.6%; Pred. No. 22;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 CIGRLCVQDGF 16
|||

DB 58 CIGRLCAQDPY 68

RESULT 12

US-10-095-450-6
; Sequence 6, Application US/10095450
; Publication No. US20020115825A1
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: A7242
; CURRENT APPLICATION NUMBER: US/10/095,450
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/127,815
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Zonulin Peptide Antagonist
US-10-095-450-6

Query Match 50.0%; Score 45; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GRLCVQDG 15
|||||

DB 1 GRLCVQDG 8

RESULT 13

US-10-648-642-6
; Sequence 6, Application US/10648642
; Publication No. US20040082512A1
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; TITLE OF INVENTION: PEPTIDE ANTAGONISTS OF ZONULIN AND METHODS FOR USE OF
; TITLE OF INVENTION: THE SAME

; FILE REFERENCE: A7242

; CURRENT APPLICATION NUMBER: US/10/648,642
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/10/095,450
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/127,815
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Zonulin Peptide Antagonist
US-10-648-642-6

Query Match 50.0%; Score 45; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GRLCVQDG 15
|||||

DB 1 GRLCVQDG 8

RESULT 14

US-10-275-877-6
; Sequence 6, Application US/10275877
; Publication No. US20050065074A1
; GENERAL INFORMATION:
; APPLICANT: PASANO, Alessio
; APPLICANT: MATTS, Tammara L.
; TITLE OF INVENTION: METHOD OF USE OF PEPTIDE ANTAGONISTS OF ZONULIN TO
; TITLE OF INVENTION: PREVENT OR DELAY THE ONSET OF DIABETES
; FILE REFERENCE: F143822
; CURRENT APPLICATION NUMBER: US/10/275,877
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 60/205,305
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-275-877-6

Query Match 50.0%; Score 45; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GRLCVQDG 15
|||||

DB 1 GRLCVQDG 8

RESULT 15

US-11-074-727A-6
; Sequence 6, Application US/11074727A
; Publication No. US20050209161A1
; GENERAL INFORMATION:
; APPLICANT: University of Maryland, Baltimore
; APPLICANT: Pasano, Alessio
; TITLE OF INVENTION: Peptide Antagonists of Zonulin and Methods for Use of the Same
; FILE REFERENCE: 006552.00102
; CURRENT APPLICATION NUMBER: US/11/074,727A
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 8
; TYPE: PRT

; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Zonulin Peptide Antagonist
 US-11-074-727A-6

Query Match 50.0%; Score 45; DB 6; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 GRLCVQDG 15
 |||||
 Db 1 GRLCVQDG 8

Search completed: January 23, 2006, 12:09:00
 Job time : 37.7511 secs

This page blank (used)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:20:59 ; Search time 76.3319 Seconds
(without alignments)
132.392 Million cell updates/sec

Title: US-10-501-838a-24
Sequence: 1 NYFLVNLAFASMAAFNTVNF 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1908s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	23	ADB16907	Adb16907 Human neu
2	114	100.0	23	ABE08241	Aeb08241 Human NK-
3	114	100.0	29	ADB16917	Adb16917 Escherich
4	114	100.0	29	ABE08251	Aeb08251 Penetrati
5	114	100.0	311	AAR28800	Har28800 Human neu
6	114	100.0	311	AAU91342	Aau91342 Neurokinin
7	114	100.0	407	AAR27524	Aar27524 Human rec
8	114	100.0	407	AAR25306	Aar25306 hepr5ults2
9	114	100.0	407	AAR33799	Aar33799 Substance
10	114	100.0	407	ABBS5366	ABBS5366 Non-endog
11	114	100.0	407	ABG76988	ABG76988 Human tac
12	114	100.0	407	ABG76987	ABG76987 Human tac
13	114	100.0	407	ABG76990	ABG76990 Human tac
14	114	100.0	407	ABG76989	ABG76989 Human tac
15	114	100.0	407	ABG77000	ABG77000 Human tac
16	114	100.0	407	AAU91345	Aau91345 Neurokinin
17	114	100.0	407	AAU91344	Aau91344 Neurokinin
18	114	100.0	407	AAU78238	Aau78238 Human Tac
19	114	100.0	407	ABBS9006	ABBS9006 Mutant hu
20	114	100.0	407	ABP81918	ABP81918 Human tac
21	114	100.0	407	ADO29656	ADO29656 Human GPC
22	109	95.6	372	AAH66623	AAH66623 Rat subst
23	109	95.6	407	AAR12970	Aar12970 Substance
24	109	95.6	407	AAR32800	Aar32800 Rat Subst

25	109	95.6	407	8	ADO29657
26	105	92.1	415	8	ABM83920
27	105	92.1	465	2	AAK41973
28	105	92.1	465	4	ABBS6368
29	105	92.1	465	6	ABP81856
30	105	92.1	465	8	ADO29660
31	105	92.1	505	5	ABG31299
32	101	88.6	452	2	ADO29661
33	100	87.7	336	2	AAK48729
34	100	87.7	336	2	AAW02701
35	100	87.7	411	4	AAE66624
36	85	74.6	440	6	ABP81685
37	83	72.8	138	8	ABO58795
38	79	69.3	23	6	ADB16894
39	79	69.3	23	9	ABE08226
40	79	69.3	30	6	ADB16914
41	79	69.3	30	9	ABE08248
42	79	69.3	301	8	ADM72122
43	79	69.3	313	4	AAU08337
44	79	69.3	313	6	ABG73541
45	79	69.3	313	6	ABU62487

ALIGNMENTS

RESULT 1
ID ADB16907 standard; peptide; 23 AA.

ADB16907;
20-NOV-2003 (first entry)
Human neurokinin 1 (NK-1) receptor penetrating peptide 29.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;
infertility; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
osteopathic; cytostatic; nootropic.

Homo sapiens.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-IB000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Gasson SA, Cohen B;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating
or preventing e.g. endocrine disorders.

Claim 2; Page 14; 60pp; English.

This invention relates to a novel peptide sequences capable of
translocating across a biological barrier. Furthermore, it refers to
methods that use these peptides to facilitate penetration of a
biologically active effector molecule such as a drug or other therapeutic
agent across biological barriers e.g. epithelial or endothelial cells
sealed by tight junctions. This peptide is derived from a bacterial
toxin, an integral membrane or extracellular protein and can comprise an
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
or enzyme. The effector molecule, however, can comprise for example
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, hematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cytoskeletal or neurotropic activities. This peptide is from the human
 CC neurokinin 1 (NK-1) receptor and is penetrating peptide 29 of the
 CC invention.

CC Sequence 23 AA;

CC Query Match 100.0%; Score 114; DB 6; Length 23;

CC Best Local Similarity 100.0%; Pred. No. 1.3e-11;

CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 1 NYFLVNLAFASMAAFNTVNF 23

CC 1 NYFLVNLAFASMAAFNTVNF 23

CC AEB08241 standard; peptide; 23 AA.

CC AEB08241;

CC 25-AUG-2005 (first entry)

CC Human NK-1 receptor penetrating peptide 29, SEQ ID NO: 24.

CC

CC Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;

CC antidiabetic; endocrine disease; gastrointestinal disease;

CC metabolic disorder; hormone deficiency; osteoporosis; osteopathic;

CC neurodegenerative disease; neuroprotective; Alzheimer's disease;

CC neurotropic; neurological disease; parkinson's disease; antiparkinsonian;

CC dementia; multiple sclerosis; immune disorder; Huntington's chorea;

CC anticonvulsant; genetic disorder; cardiovascular disease;

CC cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;

CC coronary artery disease; cardiac; vasotropic; obesity; anorectic;

CC nutritional disorder; vitamin deficiency; renal disease; nephrotropic;

CC genitourinary disease; hematological disease; antianemic; anemia;

CC autoimmune disease; immunosuppressive; immune deficiency;

CC immunostimulant; infectious disease; antimicrobial; infection;

CC erectile dysfunction; andrology; major depressive disorder;

CC antidepressant; psychiatric disorder; pain; analgesic;

CC bacterial infection; antibacterial; viral infection; virocidic;

CC fungal infection; fungicide; parasitic infection; antiparasitic;

CC renal failure; antifertility; antipneumatic; cyostatic;

CC antiinflammatory; hepatotropic; hepatitis B virus infection; receptor;

CC Natural Killer 1; NK1.

CC Homo sapiens.

CC US2005136103-A1.

CC 23-JUN-2005.

CC 16-SEP-2004; 2004US-00942300.

CC 17-SEP-2003; 2003US-00664989.

CC 17-SEP-2003; 2003US-00665184.

CC 17-SEP-2003; 2003US-0503615P.

CC (BENS/) BEN-SASSON S A.

CC (COHE/) COHEN E.

CC Ben-Sasson SA, Cohen E;

CC WPI, 2005-444089/45.

PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.

PS Claim 57; SEQ ID NO 24; 59pp; English.

CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the human NK-1 penetrating peptide.
 CC This sequence is used in the effective translocation of aminoglycoside
 CC antibiotics and antifungal agents across an epithelial barrier.

CC Sequence 23 AA;

CC Query Match 100.0%; Score 114; DB 9; Length 23;

CC Best Local Similarity 100.0%; Pred. No. 1.3e-11;

CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 1 NYFLVNLAFASMAAFNTVNF 23

CC 1 NYFLVNLAFASMAAFNTVNF 23

CC ADB16917 standard; peptide; 23 AA.

CC ADB16917;

CC 20-NOV-2003 (first entry)

CC Escherichia coli peptide 3 coupled to imaging compound linker, IBW-007.

CC epithelial; endothelial; tight junction; diabetes; infertility; hormone;

CC vitamin deficiency; neurodegenerative; cardiovascular; hematological;

CC endocrine disorder; obesity; neoplastic disease; neuroprotective;

CC cardiac; antiarteriosclerotic; osteopathic; cyostatic; neurotropic;

CC imaging linker; penetrating peptide; IBW-007.

CC Synthetic.

CC Escherichia coli.

CC Key Location/Qualifiers

CC Modified-site 1 /note= "N-terminal acetyl"

CC Modified-site 29 /note= "C-terminal amide"

CC WO2003066859-A2.

CC 14-AUG-2003.

CC 07-FEB-2003; 2003WO-IB000968.

CC 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX
 XX Example 3, Page 40; 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/macocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, hematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX Accordingly, the peptides of this invention can be used in compositions
 XX that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 XX cytostatic or neurotropic activities. This peptide sequence is IBM-007,
 XX which consists of the Escherichia coli penetrating peptide 3 coupled to
 XX the imaging linker peptide used in an exemplification of the invention.
 XX
 XX Sequence 29 AA:
 XX
 XX Query Match 100.0%; Score 114; DB 6; Length 29;
 XX Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 NYFLVNLAPAEASMAAFNTVYNP 23
 XX 1 NYFLVNLAPAEASMAAFNTVYNP 23
 XX
 XX DB 1 NYFLVNLAPAEASMAAFNTVYNP 23
 XX
 XX RESULT 4
 XX ABB08251
 XX ID ABB08251 standard; peptide; 29 AA.
 XX
 XX ABB08251;
 XX
 XX 25-ANG-2005 (first entry)
 XX
 XX Penetrating peptide SEQ: 34 used in composition for mucosal vaccination.
 XX
 XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX anti-diabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 XX coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX genitourinary disease; hematological disease; anti-anemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; virucide;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antirheumatic; cytostatic;

KM anti-inflammatory; hepatotropic; hepatitis B virus infection.
 XX unidentified.
 XX
 XX Key Location/Qualifiers
 XX Misc-difference 1
 XX FT /note= "N-terminal acylated"
 XX FT Misc-difference 29
 XX FT /note= "Optionally C-terminal amide, optionally the free
 XX FT amino group of lysine is acylated with a fatty acid"
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHR/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 63; SEQ ID NO 34; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for transepithelial delivery of effector. The
 XX invention comprises the effector sequentially coupled with a counter ion
 XX and at least one hydrophobic agent, where the effector is selectively
 XX encapsulated into a complex. The invention is useful for translocating
 XX effectors across a biological barrier such as epithelial cells and
 XX endothelial cells during treatment and prevention of disease or
 XX pathological conditions (including endocrine disorders, diabetes,
 XX infertility, hormone deficiencies, osteoporosis, ophthalmological
 XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 XX Parkinson's disease, multiple sclerosis, hyper-coagulable states, hypo-
 XX cardiovascular disorders, atherosclerosis, cerebrovascular events, metabolic
 XX coagulable states, coronary disease, renal disorders, renal failure,
 XX disorders, obesity, vitamin deficiencies, immune deficiencies and
 XX hematological disorders, autoimmune diseases, immune deficiencies,
 XX infectious diseases, viral infections, bacterial infections, fungal
 XX infections, parasitic infections, neoplastic diseases, multi-factorial
 XX disorders, impotence, chronic pain, depression, different fibrosis states
 XX and short stature) and for mucosal vaccination against anthrax and
 XX hepatitis B. The present sequence is a penetrating peptide (IBM-007) used
 XX in the composition for mucosal vaccination using a counter anion and a
 XX penetrating peptide.
 XX
 XX Sequence 29 AA:
 XX
 XX Query Match 100.0%; Score 114; DB 9; Length 29;
 XX Best Local Similarity 100.0%; Pred. No. 1.7e-11;
 XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 NYFLVNLAPAEASMAAFNTVYNP 23
 XX 1 NYFLVNLAPAEASMAAFNTVYNP 23
 XX
 XX DB 1 NYFLVNLAPAEASMAAFNTVYNP 23
 XX
 XX RESULT 5
 XX AAR28800
 XX ID AAR28800 standard; protein; 311 AA.
 XX
 XX AAR28800;

```

AC AAR28800;
XX
XX 25-MAR-2003 (revised)
DT 22-MAR-1993 (first entry)
XX
XX Human neurokinin-1 receptor.
XX
XX NK1R sf; arthritis; Substance P.
XX
XX Homo sapiens.
XX
XX EP514207-A2.
XX
XX 19-NOV-1992.
XX
XX 15-MAY-1992; 92BP-00304432.
XX
XX 17-MAY-1991; 91US-00701930.
XX
XX 17-MAY-1991; 91US-00701935.
XX
XX 17-MAY-1991; 91US-00701937.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Strader CD, Fong TM;
XX
XX WPI; 1992-384034/47.
XX
XX N-PSDB; AAQ30557.
XX
XX New human neurokinin-1 receptor short form protein - useful for
XX identifying and determining substance P antagonists in arthritic
XX patients.
XX
XX Claim 3; Fig 1; 36pp; English.
XX
XX Human neurokinin-1 receptor short form protein can be used to identify
XX and evaluate ligands which bind to the receptor. The NK1R sf protein is
XX suitable for use in an assay of substance P antagonists in arthritis
XX patients. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX
XX Sequence 311 AA;
XX
XX
XX Query Match 100.0%; Score 114; DB 2; Length 311;
XX Best Local Similarity 100.0%; Pred. No. 2.9e-10;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 NYPLVNLAFARSMARFTVNF 23
XX |||||
XX 68 NYPLVNLAFARSMARFTVNF 90
XX
XX
XX RESULT 6
XX AAU91342
XX ID AAU91342 standard; protein; 311 AA.
XX
XX AAU91342;
XX
XX 18-UN-2002 (first entry)
XX
XX Neurokinin 1 receptor (NK-1) #1.
XX
XX
XX Human; neurokinin receptor-1; NK-1; dermatological disorder;
XX immune disorder; autoimmune disorder; cardiovascular disorder;
XX vascular disorder; airway disorder; neuropathic disorder; pain;
XX psychiatric disorder; central nervous system disorder; inflammation;
XX respiratory condition; ophthalmic condition; intestinal condition;
XX demyelinating disease; small cell lung cancer; depression;
XX hypersensitivity disorder; allergy; vasospastic disease; alcoholism;
XX neurodegenerative disorder; acquired immune deficiency syndrome; AIDS;
XX neuro-pathological disorder; stress; receptor.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO200213799-A2.

```

```

XX
XX 21-FEB-2002.
XX
XX 17-AUG-2001; 2001WO-IB001510.
XX
XX 18-AUG-2000; 2000US-0226086P.
XX
XX (UVMC-) UNIV MCGILL.
XX
XX Henry JL, Cahill CM, Yaehpal K;
XX
XX WPI; 2002-241835/29.
XX
XX N-PSDB; ABK63834.
XX
XX
XX Treating pathological condition involving neurokinin receptor-1, e.g.
XX pain or inflammation, by administering oligonucleotide or a non-
XX nucleotide disruptor compound which modulate NK-1 receptor biosynthetic
XX pathway.
XX
XX Example 18; Page 60; 100pp; English.
XX
XX
XX The invention relates to a method of treating a pathological condition
XX characterised partially by involvement of neurokinin receptor-1 (NK-1)
XX receptor, especially treating, attenuating or preventing pain or
XX inflammatory condition. The method comprises administering to a mammal, a
XX compound chosen from an oligonucleotide, its analogue, and a disruptor
XX which interferes with function or production of NK-1 receptors. The
XX method is useful for treating a pathological condition characterised by
XX involvement of NK-1 receptor such as dermatological, immune, autoimmune,
XX cardiovascular, vascular disorders (e.g. migraine), airway, neuropathic,
XX psychiatric and central nervous system disorders (e.g. anxiety,
XX psychosis, schizophrenia), gut inflammation, arthritis, and central or
XX peripheral aspects of chronic or acute pain, and for treating,
XX attenuating or preventing pain or inflammation such as peripheral,
XX chronic, acute pain or inflammation, neuropathic pain, inflammation or
XX pain relating to psychiatric disorders and central nervous system
XX disorders, including hyperalgesia, allodynia, neuralgia and dysesthesia
XX in a mammal, in particular human. NK-1 receptor related disorders,
XX diseases, or pathological conditions treatable by this method include
XX respiratory conditions (e.g. asthma, allergic rhinitis), ophthalmic
XX conditions (e.g. conjunctivitis), cutaneous/dermatological conditions
XX (allergic dermatitis), intestinal conditions (e.g. ulcerative colitis and
XX Crohn's disease), cardiovascular conditions (stroke), chronic
XX gastrointestinal tract inflammation, and inflammatory diseases such as
XX inflammatory bowel diseases. Other disorders and diseases include
XX cardiovascular pathologies including stroke, chronic inflammatory
XX diseases such as rheumatoid arthritis, demyelinating diseases such as
XX multiple sclerosis, small cell lung cancer, depression, hypersensitivity
XX disorders such as allergies and poison ivy, vasospastic diseases such as
XX angina, addiction disorders such as alcoholism, neurodegenerative
XX disorders such as acquired immune deficiency syndrome (AIDS) related
XX dementia, epilepsy, Alzheimer's disease, and other neuro-pathological
XX disorders such as peripheral neuropathy, oedema, stress related and
XX somatic disorders, and osteoarthritis. Antisense oligonucleotides
XX effectively treat chronic conditions and other pathological states
XX without the co-administration of substance P, and reduce the number of
XX activated receptors while not reducing the number of quiescent NK-1
XX receptors. Receptors not chronically stimulated are less affected,
XX reducing side effects of treatment. AAU91342-AAU91346 represent NK-1
XX receptor amino acid sequences of the invention
XX
XX
XX Sequence 311 AA;
XX
XX
XX Query Match 100.0%; Score 114; DB 5; Length 311;
XX Best Local Similarity 100.0%; Pred. No. 2.9e-10;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 NYPLVNLAFARSMARFTVNF 23
XX |||||
XX 68 NYPLVNLAFARSMARFTVNF 90
XX
XX
XX RESULT 7

```

AA27524
ID AAR27524 standard; protein; 407 AA.
XX
XX
AC AAR27524;
XX
DT 25-MAR-2003 (revised)
DT 05-MAR-1993 (first entry)
XX
DE Human recombinant NK-1 receptor.
XX
XX Neurokinin; substance P; inflammation; pain; mental illness; antagonist.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH Domain 1..31 "extracellular domain"
FT /note="extracellular domain"
FT Domain 60..67 "intracellular domain"
FT /note="intracellular domain"
FT Domain 90..101 "extracellular domain"
FT /note="extracellular domain"
FT Domain 129..148 "intracellular domain"
FT /note="intracellular domain"
FT Domain 170..197 "extracellular domain"
FT /note="extracellular domain"
FT Domain 220..248 "intracellular domain"
FT /note="extracellular domain"
FT Domain 271..287 "intracellular domain"
FT /note="intracellular domain"
FT Domain 310..407 "intracellular domain"
FT /note="intracellular domain"
XX
XX MO9216547-A1.
XX
XX 01-OCT-1992.
XX
XX 13-MAR-1992; 92WO-US002007.
XX
XX 15-MAR-1991; 91US-00670039.
XX
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX
XX Gerard NP, Gerard C;
XX
XX WPI; 1992-349150/42.
XX
XX N-PSDB; AAQ29179.
XX
XX Human recombinant NK-1 receptor and fragments - for treating
XX inflammation, pain and mental illness and for screening of other cpds.
XX that antagonise interaction between substance P neuro-transmitter and NK-
XX 1 receptors.
XX
XX Claim 2; Fig 1; 27pp; English.
XX
XX The human neurokinin (NK-1) receptor sequence was deduced from the DNA
XX sequence obd. by screening a human placental genomic DNA library with a
XX probe having a partial sequence of human NK-1 receptor. The proteins
XX domains were determined by comparison with other members of the rhodopsin
XX superfamily and by examination of the hydrophobic and hydrophilic regions
XX of the mol. An extracellular domain may be capable of binding substance P
XX neurotransmitter. The NK-1 receptor may be used in therapeutic compns.
XX for the treatment of pain, inflammatory diseases e.g. arthritis or
XX asthma, mental illness e.g. schizophrenia, and stress. The compsn. is
XX administered in an amt. effective to antagonise an interaction between
XX substance P transmitter and an NK-1 receptor. See also AAQ29180-1.
XX (Updated on 25-MAR-2003 to correct FN field.)
XX
XX
XX Sequence 407 AA;
SQ

Query Match 100.0%; Score 114; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 4e-10; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFARASMAAFNTVNF 23
DB 68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 8
AAR25306
ID AAR25306 standard; protein; 407 AA.
XX
XX AAR25306;
XX
XX 25-MAR-2003 (revised)
DT 21-DEC-1998 (revised)
DT 15-MAR-1993 (first entry)
XX
XX hepsuts2 primer 86-105.
XX
XX Human; neurokinin-1 receptor; NK1R; membrane receptor; substance P;
XX neurotransmitter; polymerase chain reaction; PCR; rat NK1R; primer;
XX amplify; probe.
XX
XX Homo sapiens.
XX
XX EP510878-A1.
XX
XX 28-OCT-1992.
XX
XX 16-APR-1992; 92EP-00303457.
XX
XX 25-APR-1991; 91US-00691197.
XX 25-APR-1991; 91US-00691198.
XX 25-APR-1991; 91US-00691200.
XX
XX (MERI) MERCK & CO INC.
XX
XX Fong TM, Strader CD;
XX
XX WPI; 1992-359073/44.
XX
XX N-PSDB; AAQ29749.
XX
XX New recombinant human neurokinin-1 receptor - used to detect and evaluate
XX PT substances that bind to substance P receptor, and to determine substance
XX PT P in body fluid of arthritis patients.
XX
XX
XX Claim 1; Fig 2; 35pp; English.
XX
XX The sequence given is encoded by the human neurokinin-1 receptor (NK1R)
XX cDNA. Human NK1R is a membrane receptor for the neurotransmitter
XX substance P. The cDNA was derived using primers which were designed using
XX CC regions of the human NK1R cDNA and also regions of the rat NK1R which
XX CC were thought to be similar to human regions. See also AAQ29670-82,
XX CC AAQ29783 and AAQ29746-48. Part of the human cDNA sequence was derived by
XX CC amplification using the primers. The remaining part of human NK1R cDNA
XX CC was obtained from a human cDNA library utilising portions of PCR
XX CC generated fragments as probes. (Revised entry submitted to correct
XX CC Pco(353) to Arg(353) in the sequence.) (Updated on 25-MAR-2003 to correct
XX CC FN field.)
XX
XX
XX Sequence 407 AA;
SQ

Query Match 100.0%; Score 114; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 4e-10; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFARASMAAFNTVNF 23
DB 68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 9
AAR32799
ID AAR32799 standard; protein; 407 AA.
XX

AC AAR32799;
 XX
 DT 25-MAR-2003 (revised)
 DT 17-JUN-1993 (first entry)
 XX
 DE Substance P receptor.
 XX
 KW human substance P receptor protein; SP; neurotransmitter; neuromodulator;
 KW central nervous system; peripheral nervous system;
 KW gastrointestinal disorders; inflammation; immune disease.
 XX
 OS Synthetic.
 XX
 PN WO9303137-A1.
 XX
 PF 18-FEB-1993.
 XX
 PF 05-AUG-1992; 92WO-US006532.
 XX
 PR 07-AUG-1991; 91US-00741200.
 XX
 PA (UNITW) UNITV WASHINGTON.
 XX
 PI Krause JB;
 XX
 DE WPI; 1993-076495/09.
 DM N-PSDB; AAO37210.
 XX
 PT New human substance P receptor protein and DNA encoding it - used e.g.
 PT for screening substance P antagonists.
 PS Example; Page 19; 40pp; English.
 XX
 CC This sequence represents the human substance P receptor. This mediates
 CC many of the biological activities of substance P (SP) so cells expressing
 CC this protein are useful in studying cellular mechanisms involved in
 CC expression of SP mRNA and for screening SP antagonists (potentially
 CC useful in treatment of central and peripheral nervous system and
 CC gastrointestinal disorders, inflammation and immune disease). Transformed
 CC cells express many more receptors than natural human tissue cells, so
 CC will allow more rapid screening of antagonists. The protein may also be
 CC useful diagnostically to identify abnormal receptors associated with
 CC disease. from IM-9 CDNA. It was amplified by PCR using AAO37208.9. The
 CC amplified fragment was blunt end ligated into SmaI-digested PBS.
 CC Transforms were further analysed by restriction mapping and by
 CC sequence analysis. One CDNA was isolated after restriction with HindIII
 CC and BamHI (present in the PBS polylinker) and made blunt ended with
 CC Klenow, as was pm2 after BamHI digestion. The CDNA was ligated to pm2 and
 CC used to transform E. coli XL-1 Blue cells by electroporation. (Updated on
 CC 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 407 AA;
 XX
 Query Match 100.0%; Score 114; DB 2; Length 407;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYFLVNLAFARASMAAFNTVNF 23
 |||||
 DB 68 NYFLVNLAFARASMAAFNTVNF 90
 |||||
 RESULT 10
 ABB56366
 ID ABB56366 standard; protein; 407 AA.
 XX
 AC ABB56366;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Non-endogenous human GPCR protein, SEQ ID NO: 525.
 XX
 KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;

KW constitutively activated GPCR; agonist; disease.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200177172-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 05-APR-2001; 2001WO-US011098.
 XX
 PR 07-APR-2000; 2000US-0195747P.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Lehmann-Bruinsma K, Liauw CW, Lin I;
 XX
 DR WPI; 2001-648759/74.
 DR N-PSDB; ABI98002.
 XX
 PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
 PT disease treatment, comprises contacting candidate compounds with versions
 PT of GPCRs.
 PS Claim 1; Page 326-327; 394pp; English.
 XX
 CC The invention relates to G protein-coupled receptors (GPCRs) for which
 CC the endogenous ligand has been identified. Non-endogenous constitutively
 CC activated versions of known GPCRs are used in the invention for the
 CC direct identification of candidate compounds as receptor agonists,
 CC inverse agonists or partial agonists. Such agonists are useful as
 CC therapeutic agents for diseases or disorders associated with GPCRs. The
 CC present sequence is a non-endogenous version of a known human GPCR
 XX
 SQ Sequence 407 AA;
 XX
 Query Match 100.0%; Score 114; DB 4; Length 407;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYFLVNLAFARASMAAFNTVNF 23
 |||||
 DB 68 NYFLVNLAFARASMAAFNTVNF 90
 |||||
 RESULT 11
 ABG76988
 ID ABG76988 standard; protein; 407 AA.
 XX
 AC ABG76988;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Human tachykinin receptor 1 variant #1.
 XX
 KW Aminopeptidase P; XPNR2; bradykinin receptor B1; human; BDKRB1;
 KW tachykinin receptor B1; TACR1; C1 esterase inhibitor; CLNH; kallikrein 1;
 KW KMT1; bradykinin receptor B2; BDKRB2; gene therapy;
 KW angiotensin converting enzyme 2; ACE2; protease inhibitor 4; P14;
 KW polymorphism; haemangioma; tumour; sarcoma; Crohn's disease; trachoma;
 KW cardiovascular disease; angina pectoris; hypertension; heart failure;
 KW myocardial infarction; ventricular hypertrophy; vascular disease;
 KW aneurysm; embolism; thrombosis; coronary artery disease; angiodaema;
 KW arteriosclerosis; atherosclerosis; hypersensitivity; sepsis;
 KW autoimmune disease; inflammatory arthritis; cancer; wound;
 KW viral infection; bacterial infection; fungal infection; COPD;
 KW Chronic obstructive pulmonary disease; enterocolitis.
 XX
 OS Homo sapiens.
 XX
 PN WO200261131-A2.
 XX
 PD 08-AUG-2002.

XX 03-DEC-2001; 2001WO-US047235.
 PF
 XX
 XX 04-DEC-2000; 2000US-0251015P.
 PR 23-JAN-2001; 2001US-0263678P.
 PR 02-MAR-2001; 2001US-0273037P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA (TSUC/) TSUCHIHASHI Z.
 PA (HUI/) HUI L.
 XX Teuchihashi Z, Hui L, Zerba KE, Ma-Edmonds M, Perrone MH;
 PI Swanson BN, Powell JR;
 XX
 XX MPI: 2002-619265/66.
 DR N-PSDB; ABS59839.
 DR
 PT New isolated nucleic acid with at least one polymorphic position, useful
 PT for detecting, diagnosing and treating disorders such as angioedema,
 PT cancer, viral, bacterial or fungal infection, cardiovascular and
 PT autoimmune diseases.
 XX
 PS Disclosure; Fig 8; 977pp; English.
 XX
 XX The invention relates to an isolated nucleic acid from a human gene
 CC encoding aminopeptidase P (XPNEP2), bradykinin receptor B1 (BDKRB1),
 CC tachykinin receptor B1 (TRCR1), Cl esterase inhibitor (C1NH), kallikrein
 CC 1 (KLK1), bradykinin receptor B2 (BDKRB2), angiotensin converting enzyme
 CC 2 (ACE2), or protease inhibitor 4 (PI4), comprising at least one
 CC polymorphic position. Also included are (1) a probe that hybridizes to a
 CC polymorphic position as provided in the detailed summary of single
 CC nucleotide polymorphisms comprising additional 5' and 3' flanking genomic
 CC sequence; (2) analysing (M1) at least one nucleic acid sample comprising
 CC obtaining the sample from one or more individuals and determining the
 CC nucleic acid sequence at one or more polymorphic positions in a gene
 CC encoding a protein selected from the group above; (3) constructing (M2)
 CC haplotypes using the genes comprising grouping at least two nucleic acids
 CC; (4) identifying (M3) an individual at risk of developing a disorder
 CC upon administration of an ACE inhibitor and/or vasopeptidase inhibitor
 CC using the polymorphic data; (5) a library of nucleic acids, each of which
 CC comprises one or more polymorphic positions within a gene encoding a
 CC human protein selected from the group above; and (6) genotyping (M4) an
 CC individual comprising obtaining a nucleic acid sample, determining the
 CC nucleotide present in at least one polymorphic position, and comparing at
 CC least one position with a known data set. The genes, (M1, M2, M3 and M4)
 CC and compositions are useful for detecting, diagnosing, treating,
 CC preventing various disorders such as angioedema and diseases which
 CC involve angiogenesis like haemangiomas, tumours, sarcomas, Crohn's
 CC disease, trachomas, and cardiovascular diseases like angina pectoris,
 CC hypertension, heart failure, myocardial infarction, ventricular
 CC hypertrophy, vascular diseases, aneurysm, embolism, thrombosis, coronary
 CC artery disease, arteriosclerosis and/or atherosclerosis, and
 CC hypersensitivity reactions, sepsis, autoimmune diseases, inflammatory
 CC arthritis, cancer, wounds, viral, bacterial or fungal infection, Chronic
 CC obstructive pulmonary disease (COPD) and enterocolitis (many other
 CC diseases and disorders are listed in the specification). The
 CC polynucleotides are also useful for chromosome identification. Antibodies
 CC against the proteins may be utilised for immunophenotyping of cell lines
 CC and biological samples. The present sequence represents a polymorphic
 CC variant of one of the proteins listed above
 XX
 SO Sequence 407 AA:
 Query Match 100.0%; Score 114; DB 5; Length 407;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ABG76987
 ID ABG76987 standard; protein; 407 AA.
 XX
 XX
 AC ABG76987;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 XX Human tachykinin receptor 1.
 XX
 XX Aminopeptidase P; XPNEP2; bradykinin receptor B1; human; BDKRB1;
 KW tachykinin receptor B1; TRCR1; Cl esterase inhibitor; C1NH; kallikrein 1;
 KW KLK1; bradykinin receptor B2; BDKRB2; gene therapy;
 KW angiotensin converting enzyme 2; ACE2; protease inhibitor 4; PI4;
 KW polymorphism; haemangioma; tumour; sarcoma; Crohn's disease; trachoma;
 KW cardiovascular disease; angina pectoris; hypertension; heart failure;
 KW myocardial infarction; ventricular hypertrophy; vascular disease;
 KW aneurysm; embolism; thrombosis; coronary artery disease; angioedema;
 KW arteriosclerosis; atherosclerosis; hypersensitivity; sepsis;
 KW autoimmune disease; inflammatory arthritis; cancer; wound;
 KW viral infection; bacterial infection; fungal infection; COPD;
 KW Chronic obstructive pulmonary disease; enterocolitis.
 XX
 OS Homo sapiens.
 XX
 XX WO200261131-A2.
 XX
 XX 08-AUG-2002.
 XX
 XX 03-DEC-2001; 2001WO-US047235.
 XX
 XX 04-DEC-2000; 2000US-0251015P.
 XX PR 23-JAN-2001; 2001US-0263678P.
 XX PR 02-MAR-2001; 2001US-0273037P.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA (TSUC/) TSUCHIHASHI Z.
 PA (HUI/) HUI L.
 XX Teuchihashi Z, Hui L, Zerba KE, Ma-Edmonds M, Perrone MH;
 PI Swanson BN, Powell JR;
 XX
 XX MPI: 2002-619265/66.
 DR N-PSDB; ABS59839.
 DR
 PT New isolated nucleic acid with at least one polymorphic position, useful
 PT for detecting, diagnosing and treating disorders such as angioedema,
 PT cancer, viral, bacterial or fungal infection, cardiovascular and
 PT autoimmune diseases.
 XX
 PS Disclosure; Fig 7; 977pp; English.
 XX
 XX The invention relates to an isolated nucleic acid from a human gene
 CC encoding aminopeptidase P (XPNEP2), bradykinin receptor B1 (BDKRB1),
 CC tachykinin receptor B1 (TRCR1), Cl esterase inhibitor (C1NH), kallikrein
 CC 1 (KLK1), bradykinin receptor B2 (BDKRB2), angiotensin converting enzyme
 CC 2 (ACE2), or protease inhibitor 4 (PI4), comprising at least one
 CC polymorphic position. Also included are (1) a probe that hybridizes to a
 CC polymorphic position as provided in the detailed summary of single
 CC nucleotide polymorphisms comprising additional 5' and 3' flanking genomic
 CC sequence; (2) analysing (M1) at least one nucleic acid sample comprising
 CC obtaining the sample from one or more individuals and determining the
 CC nucleic acid sequence at one or more polymorphic positions in a gene
 CC encoding a protein selected from the group above; (3) constructing (M2)
 CC haplotypes using the genes comprising grouping at least two nucleic acids
 CC; (4) identifying (M3) an individual at risk of developing a disorder
 CC upon administration of an ACE inhibitor and/or vasopeptidase inhibitor
 CC using the polymorphic data; (5) a library of nucleic acids, each of which
 CC comprises one or more polymorphic positions within a gene encoding a
 CC human protein selected from the group above; and (6) genotyping (M4) an
 CC individual comprising obtaining a nucleic acid sample, determining the
 CC nucleotide present in at least one polymorphic position, and comparing at
 CC least one position with a known data set. The genes, (M1, M2, M3 and M4)
 CC and compositions are useful for detecting, diagnosing, treating,

CC preventing various disorders such as angioedema and diseases which
 CC involve angiogenesis like haemangiomas, tumours, sarcomas, Crohn's
 CC disease, trachomas, and cardiovascular diseases like angina pectoris,
 CC hypertension, heart failure, myocardial infarction, ventricular
 CC hypertrophy, vascular diseases, aneurysm, embolism, thrombosis, coronary
 CC artery disease, arteriosclerosis and/or atherosclerosis, and
 CC hypersensitivity reactions, sepsis, autoimmune diseases, inflammatory
 CC arthritis, cancer, wounds, viral, bacterial or fungal infection, chronic
 CC obstructive pulmonary disease (COPD) and enterocolitis (many other
 CC diseases and disorders are listed in the specification). The
 CC polynucleotides are also useful for chromosome identification. Antibodies
 CC against the proteins may be utilised for immunophenotyping of cell lines
 CC and biological samples. The present sequence represents a polymorphic
 CC variant of one of the proteins listed above

XX
 CC Sequence 407 AA;

XX
 CC Query Match 100.0%; Score 114; DB 5; Length 407;
 CC Best Local Similarity 100.0%; Pred. No. 4e-10;
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFAEASMAAFTVVNF 23
 |||||
 DB 68 NYFLVNLAFAEASMAAFTVVNF 90

RESULT 13
 ABG76990 standard; protein; 407 AA.

XX
 AC ABG76990;

XX
 DT 05-NOV-2002 (first entry)

XX
 DE Human tachykinin receptor 1 variant #3.

XX
 XX Amino-peptidase P; XPNP2; bradykinin receptor B1; human; BDKRB1;
 KW tachykinin receptor B1; TACR1; Cl esterase inhibitor; C1NH; kallikrein 1;
 KW KLK1; bradykinin receptor B2; BDKRB2; gene therapy;
 KW angiotensin converting enzyme 2; ACE2; protease inhibitor 4; P14;
 KW polymorphisms; haemangioma; tumour; sarcoma; Crohn's disease; trachoma;
 KW cardiovascular disease; angina pectoris; hypertension; heart failure;
 KW myocardial infarction; ventricular hypertrophy; vascular disease;
 KW aneurysm; embolism; thrombosis; coronary artery disease; angioedema;
 KW arteriosclerosis; atherosclerosis; hypersensitivity; sepsis;
 KW autoimmune disease; inflammatory arthritis; cancer; wound;
 KW viral infection; bacterial infection; fungal infection; COPD;
 KW Chronic obstructive pulmonary disease; enterocolitis.

XX
 OS Homo sapiens.

XX
 XX WO200261131-A2.

XX
 PD 08-AUG-2002.

XX
 PF 03-DEC-2001; 2001WO-US047235.

XX
 PR 04-DEC-2000; 2000US-0251015P.

XX
 PR 23-JAN-2001; 2001US-0263678P.

XX
 PR 02-MAR-2001; 2001US-0273037P.

XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX
 PA (TSUC) TSUCHIHASHI Z.

XX
 PA (HUTL) HUT L.

XX
 PI Tsuchihashi Z, Hui L, Zerba KE, Ma-Edmonds M, Perrone MH;
 PI Swanson BN, Powell JR;

XX
 DR MPI. 2002-619265/66.

XX
 DR N-PsDB; ABSS9842.

PT New isolated nucleic acid with at least one polymorphic position, useful
 PT for detecting, diagnosing and treating disorders such as angioedema,

PT cancer, viral, bacterial or fungal infection, cardiovascular and
 PT autoimmune diseases.

XX
 XX Disclosure; Fig 10; 977pp; English.

XX
 XX The invention relates to an isolated nucleic acid from a human gene
 CC encoding aminopeptidase P (XPNP2), bradykinin receptor B1 (BDKRB1),
 CC tachykinin receptor B1 (TACR1), Cl esterase inhibitor (C1NH), kallikrein
 CC 1 (KLK1), bradykinin receptor B2 (BDKRB2), angiotensin converting enzyme
 CC 2 (ACE2) or protease inhibitor 4 (P14), comprising at least one
 CC polymorphic position. Also included are (1) a probe that hybridises to a
 CC polymorphic position as provided in the detailed summary of single
 CC nucleotide polymorphisms comprising additional 5' and 3' flanking genomic
 CC sequence; (2) analysing (M1) at least one nucleic acid sample comprising
 CC obtaining the sample from one or more individuals and determining the
 CC nucleic acid sequence at one or more polymorphic positions in a gene
 CC encoding a protein selected from the group above; (3) constructing (M2)
 CC haplotypes using the genes comprising grouping at least two nucleic acids
 CC ; (4) identifying (M3) an individual at risk of developing a disorder
 CC upon administration of an ACE inhibitor and/or vasopressinase inhibitor
 CC using the polymorphic data; (5) a library of nucleic acids, each of which
 CC comprises one or more polymorphic positions within a gene encoding a
 CC human protein selected from the group above; and (6) genotyping (M4) an
 CC individual comprising obtaining a nucleic acid sample, determining the
 CC nucleotide present in at least one polymorphic position, and comparing at
 CC least one position with a known data set. The genes, (M1, M2, M3 and M4)
 CC and compositions are useful for detecting, diagnosing, treating,
 CC preventing various disorders such as angioedema and diseases which
 CC involve angiogenesis like haemangiomas, tumours, sarcomas, Crohn's
 CC disease, trachomas, and cardiovascular diseases like angina pectoris,
 CC hypertension, heart failure, myocardial infarction, ventricular
 CC hypertrophy, vascular diseases, aneurysm, embolism, thrombosis, coronary
 CC artery disease, arteriosclerosis and/or atherosclerosis, and
 CC hypersensitivity reactions, sepsis, autoimmune diseases, inflammatory
 CC arthritis, cancer, wounds, viral, bacterial or fungal infection, Chronic
 CC obstructive pulmonary disease (COPD) and enterocolitis (many other
 CC diseases and disorders are listed in the specification). The
 CC polynucleotides are also useful for chromosome identification. Antibodies
 CC against the proteins may be utilised for immunophenotyping of cell lines
 CC and biological samples. The present sequence represents a polymorphic
 CC variant of one of the proteins listed above

XX
 XX Sequence 407 AA;

XX
 CC Query Match 100.0%; Score 114; DB 5; Length 407;
 CC Best Local Similarity 100.0%; Pred. No. 4e-10;
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFAEASMAAFTVVNF 23
 |||||
 DB 68 NYFLVNLAFAEASMAAFTVVNF 90

RESULT 14
 ABG76989 standard; protein; 407 AA.

XX
 AC ABG76989;

XX
 DT 05-NOV-2002 (first entry)

XX
 DE Human tachykinin receptor 1 variant #2.

XX
 XX Amino-peptidase P; XPNP2; bradykinin receptor B1; human; BDKRB1;
 KW tachykinin receptor B1; TACR1; Cl esterase inhibitor; C1NH; kallikrein 1;
 KW KLK1; bradykinin receptor B2; BDKRB2; gene therapy;
 KW angiotensin converting enzyme 2; ACE2; protease inhibitor 4; P14;
 KW polymorphisms; haemangioma; tumour; sarcoma; Crohn's disease; trachoma;
 KW cardiovascular disease; angina pectoris; hypertension; heart failure;
 KW myocardial infarction; ventricular hypertrophy; vascular disease;
 KW aneurysm; embolism; thrombosis; coronary artery disease; angioedema;
 KW arteriosclerosis; atherosclerosis; hypersensitivity; sepsis;
 KW autoimmune disease; inflammatory arthritis; cancer; wound;

KW viral infection; bacterial infection; fungal infection; COPD;
 KM Chronic obstructive pulmonary disease; enterocolitis.
 XX
 XX Homo sapiens.
 OS
 XX
 XX MO200261131-A2.
 PN
 XX
 XX 08-AUG-2002.
 XX
 PD
 XX
 PF 03-DEC-2001; 2001WO-US047235.
 XX
 PR 04-DEC-2000; 2000US-0251015P.
 XX 23-JAN-2001; 2001US-0263678P.
 PR 02-MAR-2001; 2001US-0273037P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA (TSUC/) TSUCHIHASHI Z.
 PA (HUIL/) HUI L.
 XX
 PI Tsuchihashi Z, Hui L, Zerba KE, Ma-Edmonds M, Perrone MH;
 PI Swanson BN, Powell JR;
 DR N-PADB; A8559841.
 DR
 XX
 XX WPI; 2002-619265/66.
 PT
 PT New isolated nucleic acid with at least one polymorphic position, useful
 PT for detecting, diagnosing and treating disorders such as angioedema,
 PT cancer, viral, bacterial or fungal infection, cardiovascular and
 PT autoimmune diseases.
 PS
 PS Disclosure; Fig 9; 977pp; English.
 XX
 XX The invention relates to an isolated nucleic acid from a human gene
 CC encoding aminopeptidase P (KPNP2), bradykinin receptor B1 (BDKRB1),
 CC lachrykinin receptor B1 (TRACR1), C1 esterase inhibitor (C1NH), kallikrein
 CC 1 (KLK1), bradykinin receptor B2 (BDKRB2), angiotensin converting enzyme
 CC 2 (ACE2) or proteinase inhibitor 4 (PI4), comprising at least one
 CC polymorphic position. Also included are (1) a probe that hybridises to a
 CC polymorphic position as provided in the detailed summary of single
 CC nucleotide polymorphisms comprising additional 5' and 3' flanking genomic
 CC sequence; (2) analysing (M1) at least one nucleic acid sample comprising
 CC the sample from one or more individuals and determining the
 CC nucleic acid sequence at one or more polymorphic positions in a gene
 CC encoding a protein selected from the group above; (3) constructing (M2)
 CC haplotypes using the genes comprising grouping at least two nucleic acids
 CC / (4) identifying (M3) an individual at risk of developing a disorder
 CC upon administration of an ACE inhibitor and/or vasopeptidase inhibitor
 CC using the polymorphic data; (5) a library of nucleic acids, each of which
 CC comprises one or more polymorphic positions within a gene encoding a
 CC human protein selected from the group above; and (6) genotyping (M4) an
 CC individual comprising obtaining a nucleic acid sample, determining the
 CC nucleotide present in at least one polymorphic position, and comparing at
 CC least one position with a known data set. The genes, (M1, M2, M3 and M4)
 CC and compositions are useful for detecting, diagnosing, treating,
 CC preventing various disorders such as angioedema and diseases which
 CC involve angiogenesis like haemangiomas, tumours, sarcomas, Crohn's
 CC disease, trichomas, and cardiovascular diseases like angina pectoris,
 CC hypertension, heart failure, myocardial infarction, ventricular
 CC hypertrophy, vascular diseases, aneurysm, embolism, thrombosis, coronary
 CC artery disease, arteriosclerosis and/or atherosclerosis, and
 CC hypersensitivity reactions, sepsis, autoimmune diseases, inflammatory
 CC arthritis, cancer, wounds, viral, bacterial or fungal infection, Chronic
 CC obstructive pulmonary disease (COPD) and enterocolitis (many other
 CC diseases and disorders are listed in the specification). The
 CC polymorphisms are also useful for chromosome identification. Antibodies
 CC against the proteins may be utilised for immunophenotyping of cell lines
 CC and biological samples. The present sequence represents a polymorphic
 CC variant of one of the proteins listed above
 XX
 XX Sequence 407 AA;
 SQ

[illegible]

CC upon administration of an ACE inhibitor and/or vasopeptidase inhibitor
 CC using the polymorphic data; (5) a library of nucleic acids, each of which
 CC comprises one or more polymorphic positions within a gene encoding a
 CC human protein selected from the group above; and (6) genotyping (M4) an
 CC individual comprising obtaining a nucleic acid sample, determining the
 CC nucleotide present in at least one polymorphic position, and comparing at
 CC least one position with a known data set. The genes, (M1, M2, M3 and M4)
 CC and compositions are useful for detecting, diagnosing, treating,
 CC preventing various disorders such as angioedema and diseases which
 CC involve angiogenesis like haemangiomas, tumours, sarcomas, Crohn's
 CC disease, trachomas, and cardiovascular diseases like angina pectoris,
 CC hypertension, heart failure, myocardial infarction, ventricular
 CC hypertrophy, vascular diseases, aneurysm, embolism, thrombosis, coronary
 CC artery disease, arteriosclerosis and/or atherosclerosis, and
 CC hypersensitivity reactions, sepsis, autoimmune diseases, inflammatory
 CC arthritis, cancer, wounds, viral, bacterial or fungal infection, Chronic
 CC obstructive pulmonary disease (COPD) and enterocolitis (many other
 CC diseases and disorders are listed in the specification). The
 CC polynucleotides are also useful for chromosome identification. Antibodies
 CC against the proteins may be utilised for immunophenotyping of cell lines
 CC and biological samples. The present sequence represents polymorphic
 CC variants of one of the proteins listed above
 CC
 XX
 SQ Sequence 407 AA;

Query Match 100.0%; Score 114; DB 5; Length 407;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAPAEASMAAFNTVVNF 23
 |||||
 DB 68 NYFLVNLAPAEASMAAFNTVVNF 90

Search completed: January 23, 2006, 10:28:33
 Job time : 77.3319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:32:56 ; Search time 12.5546 seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-24
Perfect score: 114
Sequence: 1 NYFLVNLAFAPASMAAFNTVNF 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	407	1 JQ1274	neurokinin 1 recep
2	114	100.0	407	2 S23510	neurokinin 1 recep
3	109	95.6	407	2 S20304	neurokinin 1 recep
4	109	95.6	407	2 A34357	neurokinin 1 recep
5	105	92.1	465	1 JQ1517	neurokinin 3 recep
6	101	88.6	385	2 S55524	neurokinin 3 recep
7	100	87.7	452	2 A34916	neurokinin 3 recep
8	85	74.6	440	2 A44081	kappa-type opioid
9	79	69.3	384	1 S00516	neurokinin 2 recep
10	79	69.3	398	2 JQ1059	neurokinin 2 recep
11	79	69.3	402	1 I56595	neurokinin 2 recep
12	78	68.4	384	2 I57957	neurokinin 2 recep
13	78	68.4	384	2 S20303	neurokinin 2 recep
14	78	68.4	390	2 A36737	neurokinin 2 recep
15	65	57.0	391	2 T32517	hypothetical prote
16	65	57.0	504	2 A41783	hypothetical prote
17	62	54.4	519	2 S17783	hypothetical prote
18	60	52.6	355	2 A42347	hypothetical prote
19	60	52.6	355	2 I51319	hypothetical prote
20	59	51.8	374	2 S28285	hypothetical prote
21	53	46.5	349	2 B45229	opsin, green-sensi
22	51	44.7	501	2 JH0447	alpha-1A-adrenergi
23	51	44.7	572	2 I39369	opsin, blue-sensi
24	50	43.9	351	2 B45229	opsin, blue-sensi
25	49	43.0	349	2 A45229	opsin, green-sensi
26	48	42.1	560	2 A38731	alpha-1A adrenergi
27	47	41.2	332	2 JCI229	adenosine receptor
28	47	41.2	332	2 I48953	adenosine receptor
29	47	41.2	332	2 A42171	A2-adenosine recep

30	47	41.2	455	1 000CG
31	47	41.2	501	2 T18863
32	46	40.4	349	1 UC5490
33	46	40.4	387	2 S55550
34	46	40.4	406	2 S55549
35	46	40.4	515	2 A40491
36	46	40.4	517	2 JCI525
37	46	40.4	517	2 A45121
38	45	39.5	123	2 F88108
39	45	39.5	345	2 T24659
40	45	39.5	353	1 JN0120
41	45	39.5	379	2 JCI5178
42	45	39.5	382	2 C86249
43	44.5	39.0	355	2 S39028
44	44.5	39.0	365	2 B46191
45	44	38.6	299	2 C97118

ALIGNMENTS

RESULT 1
JQ1274
neurokinin 1 receptor - human
N/Alternate names: NK-1 receptor, substance P receptor
C/Species: Homo sapiens (man)
C/Date: 20-Apr-2000 #sequence, revision 20-Apr-2000 #text, change 09-Jul-2004
C/Accession: A41134, JQ1274, JH0478, S21188
R/Gerard, N.P.; Garraway, L.A.; Eddy Jr., R.L.; Shows, T.B.; Iijima, H.; Paquet, J.L.;
Biochemistry 30, 10640-10646, 1991
A/Title: Human substance P receptor (NK-1) organization of the gene, chromosome locally,
A/Reference number: A41134, MIM:92031510, PMID:1657150
A/Accession: A41134
A/Molecule type: DNA
A/Residues: 1-328, G', 329-332, 334-407 <GER>
A/Cross-references: UNIPROT:P5103; UNIPARC:UPI000017351E; GB:M76675; NID:g189231
A/Note: In the authors' translation 333-Gly is shown before residue 329 and, consequently,
R/Takeda, Y.; Chou, K.B.; Takeda, J.; Sechats, B.S.; Krause, J.B.
Biochem. Biophys. Res. Commun. 179, 1232-1240, 1991
A/Title: Molecular cloning, structural characterization and functional expression of th
A/Reference number: JQ1274, MIM:92028856, PMID:1718267
A/Accession: JQ1274
A/Molecule type: mRNA
A/Residues: 1-407 <TKX>
A/Cross-references: UNIPARC:UPI0000039D68; GB:M74290; NID:g338612; PIND:AAA60601.1; PID
R/Hopkins, B.; Powell, S.J.; Danks, P.; Briggse, I.; Graham, A.
Biochem. Biophys. Res. Commun. 180, 1110-1117, 1991
A/Title: Isolation and characterization of the human lung NK-1 receptor cDNA.
A/Reference number: JH0478, MIM:92062052, PMID:1659396
A/Accession: JH0478
A/Molecule type: mRNA
A/Residues: 1-407 <HOP>
A/Cross-references: UNIPARC:UPI0000039D68; GB:S62045; NID:g237994; PIND:AAB20168.1; PID
A/Experimental source: lung
A/Note: the authors translated the codon CAA for residue 31 as Glu
R/Takanaishi, K.; Tanaka, A.; Hara, M.; Nakatani, S.
Eur. J. Biochem. 204, 1025-1033, 1992
A/Title: The primary structure and gene organization of human substance P and neuromedi
A/Reference number: S21188, MIM:92201186, PMID:1312928
A/Accession: S21188
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-407 <TKX2>
A/Cross-references: UNIPARC:UPI0000039D68; GB:X65177; NID:g36636; PIND:CAA46292.1; PID:
C/Comment: The endogenous ligand of this receptor is neurokinin 1 (substance P), one of
C/Genetics:
A/Gene: GDB:TKX1R
A/Cross-references: GDB:128977; OMIM:162323
A/Map position: 2pter-2qter
C/Superfamily: neurokinin 1 receptor
C/Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane P
P:33-58/Domain: transmembrane #status predicted <TM1>
P:69-92/Domain: transmembrane #status predicted <TM2>

F.110-128/Domain: transmembrane #status predicted <TM3>
F.149-168/Domain: transmembrane #status predicted <TM4>
F.193-221/Domain: transmembrane #status predicted <TM5>
F.249-273/Domain: transmembrane #status predicted <TM6>
F.286-308/Domain: transmembrane #status predicted <TM7>
F.14/18/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.105-180/Distal fluoride bonds: #status predicted

Query Match 100.0%; Score 114; DB 1; Length 407;
Best Local Similarity 100.0%; Pred. No. 8.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFARASMAAFNTVNF 23
|||||
Db 68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 2
S23510

neurokinin 1 receptor - guinea pig

N.Alternate names: NK-1 receptor; substance P receptor

C.Species: Cavia porcellus (guinea pig)

C.Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004

C.Accession: S23510; S19198

R.Gorbulov, V.; Akhundova, A.; Luzius, H.; Fahrenholz, F.

Biochim. Biophys. Acta 1131, 99-102, 1992

A.Title: Molecular cloning of substance P receptor cDNA from guinea-pig uterus.

A.Reference number: S23510; MUID:92256498; PMID:1374648

A.Accession: S23510

A.Molecule type: mRNA

A.Cross-references: UNIPROT:P30547; UNIPARC:UPI0000001112; EMBL:X64323; NID:g49565; PIDN

C.Superfamily: neurokinin 1 receptor

C.Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F.32-55/Domain: transmembrane #status predicted <TM1>

F.69-89/Domain: transmembrane #status predicted <TM2>

F.117-128/Domain: transmembrane #status predicted <TM3>

F.149-169/Domain: transmembrane #status predicted <TM4>

F.196-217/Domain: transmembrane #status predicted <TM5>

F.250-280/Domain: transmembrane #status predicted <TM6>

F.289-308/Domain: transmembrane #status predicted <TM7>

Query Match 100.0%; Score 114; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 8.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFARASMAAFNTVNF 23
|||||
Db 68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 3
S20304

neurokinin 1 receptor - mouse

N.Alternate names: NK-1 receptor; substance P receptor

C.Species: Mus musculus (house mouse)

C.Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C.Accession: S20304; I56216; T73044

R.Sundelin, J.B.; Froydeni, D.M.; Wahlstedt, C.R.; Laurell, H.; Pohl, J.S.; Peterson, Eur. J. Biochem. 203, 625-631, 1992

A.Title: Molecular cloning of the murine substance K and substance P receptor genes.

A.Reference number: S20303; MUID:92137253; PMID:1370937

A.Accession: S20304

A.Molecule type: preliminary

A.Status: preliminary

A.Molecule type: mRNA

A.Cross-references: UNIPROT:P30548; UNIPARC:UPI0000029850; GB:X62934; NID:g54206; PIDN:Q

R.Cook, G.A.; Elliott, D.; Metwalli, A.; Blum, A.M.; Sandor, M.; Lynch, R.; Weinstock, J.

J. Immunol. 152, 1830-1835, 1994

A.Title: Molecular evidence that granuloma T lymphocytes in murine schistosomiasis manac

A.Accession: S20304

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 63-290 <COO1>

A.Cross-references: UNIPARC:UPI000016CF5F; GB:L27826; NID:g450286; PIDN:AAA17891.1; PID:

A.Experimental source: tissue brain

A.Accession: I73044

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 63-290 <COO2>

A.Cross-references: UNIPARC:UPI000016CF5F; GB:L27826; NID:g450290; PIDN:AAA17892.1; PID:

A.Experimental source: tissue granuloma

C.Superfamily: neurokinin 1 receptor

C.Keywords: G protein-coupled receptor; transmembrane protein

Query Match 95.6%; Score 109; DB 2; Length 407;
Best Local Similarity 95.7%; Pred. No. 5.8e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFARASMAAFNTVNF 23
|||||
Db 68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 4
A34357

neurokinin 1 receptor - rat

N.Alternate names: NK-1 receptor; substance P receptor

C.Species: Rattus norvegicus (Norway rat)

C.Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004

C.Accession: A38692; A34357; A40089

R.Hershey, A.D.; Dykema, P.E.; Krause, J.E.

J. Biol. Chem. 266, 4366-4374, 1991

A.Title: Organization, structure, and expression of the gene encoding the rat substance

A.Reference number: A38692; MUID:91154239; PMID:1705552

A.Accession: A38692

A.Molecule type: DNA

A.Residues: 1-407 <HER>

A.Cross-references: UNIPROT:P14600; UNIPARC:UPI00001301E4; GB:M34751

R.Yokota, Y.; Sasai, Y.; Tanaka, K.; Fujiwara, T.; Tsuchida, K.; Shigemoto, R.; Kakizuka

J. Biol. Chem. 264, 17649-17652, 1989

A.Title: Molecular characterization of a functional cDNA for rat substance P receptor.

A.Reference number: A34357; MUID:90036822; PMID:2478537

A.Accession: A34357

A.Molecule type: mRNA

A.Residues: 1-407 <YOK>

A.Cross-references: UNIPARC:UPI00001301E4; GB:J05097; NID:g207051; PIDN:AAA42175.1; PID:

R.Hershey, A.D.; Krause, J.E.

Science 247, 958-962, 1990

A.Title: Molecular characterization of a functional cDNA encoding the rat substance P re

C.Superfamily: neurokinin 1 receptor

C.Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 95.6%; Score 109; DB 2; Length 407;
Best Local Similarity 95.7%; Pred. No. 5.8e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFARASMAAFNTVNF 23
|||||
Db 68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 5
J01517

neurokinin 3 receptor - human

N.Alternate names: neuromedin K receptor; NK-3 receptor

C.Species: Homo sapiens (man)

C.Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C.Accession: J01517; S20435; S21237

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C/Accession: S00516
 R/Masu, Y.; Nakayama, K.; Tamaki, H.; Harada, Y.; Kuno, M.; Nakamishi, S.
 Nature 329, 836-838, 1987
 A>Title: cDNA cloning of bovine substance-K receptor through oocyte expression system.
 A/Reference number: S00516; MUID:88039072; PMID:2823146
 A/Accession: S00516
 A/Molecule type: mRNA
 A/Residues: 1-384 <MAS>
 A/Cross-references: UNIPROT:P05363; UNIPARC:UPI00001301EE; EMBL:X06295; NID:g746; PIDN:C
 C/Superfamily: neurokinin 1 receptor
 C/Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane pr
 F/33-59/Domain: transmembrane #status predicted <TM1>
 F/70-93/Domain: transmembrane #status predicted <TM2>
 F/111-129/Domain: transmembrane #status predicted <TM3>
 F/150-169/Domain: transmembrane #status predicted <TM4>
 F/193-222/Domain: transmembrane #status predicted <TM5>
 F/252-275/Domain: transmembrane #status predicted <TM6>
 F/288-310/Domain: transmembrane #status predicted <TM7>
 F/11.19/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F/106-181/Disulfide bonds: #status predicted

Query Match 69.3%; Score 79; DB 1; Length 384;
 Best Local Similarity 65.2%; Pred. No. 4.3e-05;
 Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Db 1 NYFLVNLAFASMAAFNTVNF 23
 69 NYFLVNLADLCMAAFNAFNF 91

RESULT 10
 J01059
 neurokinin 2 receptor - human
 N/Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
 C/Species: Homo sapiens (man)
 C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
 C/Accession: J01059; J05075; A23658; A61224
 R/Graham, A.; Hopkins, B.; Powell, S.J.; Danks, P.; Briggs, I.
 Biochem. Biophys. Res. Commun. 177, 8-16, 1991
 A>Title: Isolation and characterization of the human lung NK-2 receptor gene using rapid
 A/Reference number: J01059; MUID:91254341; PMID:1710456
 A/Accession: J01059
 A/Molecule type: DNA
 A/Residues: 1-398 <GRA1>
 A/Cross-references: UNIPROT:P21452; UNIPARC:UPI000016ADE7; GB:M75105; NID:G189219; PIDN:
 A/Accession: J05075
 A/Molecule type: mRNA
 A/Residues: 1-398 <GRA2>
 A/Cross-references: UNIPARC:UPI000016ADE7; GB:M75105; NID:G189219; PIDN:AB05897.1; PID:
 A/Experimental source: lung
 A/Note: 23-ntle and 375 nt were also found
 R/Gerard, N.P.; Eddy Jr., R.L.; Shown, T.B.; Gerard, C.
 J. Biol. Chem. 265, 20455-20462, 1990
 A>Title: The human neurokinin A (substance K) receptor. Molecular cloning of the gene, c
 A/Reference number: A23658; MUID:91056095; PMID:2173708
 A/Accession: A23658
 A/Molecule type: DNA
 A/Residues: 1-22, 'T', 24-240, 'L', 242-398 <GER>
 A/Cross-references: UNIPARC:UPI000017351F; GB:M60284; GB:J05680; NID:G189140
 A/Note: the authors translated the codon GGA for residue 317 as Glu
 R/Kris, R.M.; South, V.; Salzman, A.; Felder, S.; Rice, G.A.; Jaye, M.; Huebner, K.; K
 Cell Growth Differ. 2, 15-22, 1991
 A>Title: Cloning and expression of the human substance K receptor and analysis of its rc
 A/Reference number: A61224; MUID:91175483; PMID:1848773
 A/Accession: A61224
 A/Molecule type: mRNA
 A/Status: not compared with conceptual translation
 A/Residues: 1-22, 'T', 24-293, 'F', 294-374, 'H', 376-398 <KRI>
 A/Cross-references: UNIPARC:UPI0000173520
 C/Comment: The endogenous ligand of this receptor is neurokinin 2 (substance K or neurok
 C/Genetics:
 A/Gene: GDB:TAC2R

A/Cross-references: GDB:126367; OMIM:162321
 A/Map position: 10q11-10q21
 A/Introns: 131/2; 196/2; 247/3; 313/2
 C/Superfamily: neurokinin 1 receptor
 C/Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane pr
 F/33-59/Domain: transmembrane #status predicted <TM1>
 F/70-93/Domain: transmembrane #status predicted <TM2>
 F/111-129/Domain: transmembrane #status predicted <TM3>
 F/150-169/Domain: transmembrane #status predicted <TM4>
 F/194-222/Domain: transmembrane #status predicted <TM5>
 F/252-275/Domain: transmembrane #status predicted <TM6>
 F/288-310/Domain: transmembrane #status predicted <TM7>
 F/11.19/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F/106-181/Disulfide bonds: #status predicted

Query Match 69.3%; Score 79; DB 1; Length 398;
 Best Local Similarity 65.2%; Pred. No. 4.5e-05;
 Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Db 1 NYFLVNLAFASMAAFNTVNF 23
 69 NYFLVNLADLCMAAFNAFNF 91

RESULT 11
 I56595
 neurokinin 2 receptor - guinea pig
 C/Species: Cavia porcellus (guinea pig)
 C/Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
 C/Accession: I56595
 R/Aharony, D.; Little, J.; Thomas, C.; Powell, S.; Downey-Jones, M.; Graham, A.
 J. Recept. Res. 14, 399-421, 1994
 A>Title: Isolation and characterization of neurokinin A receptor cDNAs from guinea-pig 1
 A/Reference number: I56595; MUID:95182423; PMID:7877137
 A/Accession: I56595
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-402 <RRS>
 A/Cross-references: UNIPROT:Q64077; UNIPARC:UPI00001301EF; GB:S76253; NID:G913274; PIDN:
 C/Superfamily: neurokinin 1 receptor

Query Match 69.3%; Score 79; DB 2; Length 402;
 Best Local Similarity 65.2%; Pred. No. 4.5e-05;
 Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Db 1 NYFLVNLAFASMAAFNTVNF 23
 69 NYFLVNLADLCMAAFNAFNF 91

RESULT 12
 I57957
 neurokinin 2 receptor - hamster
 N/Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
 C/Species: Cricetinae gen. sp. (hamster)
 C/Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 19-May-2000
 C/Accession: I57957
 R/Aharony, D.; Little, J.; Thomas, C.; Powell, S.; Berry, D.; Graham, A.
 Mol. Pharmacol. 45, 9-19, 1994
 A>Title: Isolation and pharmacological characterization of a hamster urinary bladder neu
 A/Reference number: I57957; MUID:94134065; PMID:8302285
 A/Accession: I57957
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-384 <RRS>
 A/Cross-references: UNIPARC:UPI00001301F1; GB:S68899; NID:G545230; PIDN:AC60680.1; PID:
 C/Superfamily: neurokinin 1 receptor

Query Match 68.4%; Score 78; DB 2; Length 384;
 Best Local Similarity 60.9%; Pred. No. 6.3e-05;
 Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Db 1 NYFLVNLAFASMAAFNTVNF 23

Db 69 NYPIINLALADLCMAAFNATFNF 91

RESULT 13

S20303
neurokinin 2 receptor - mouse
N/Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
C/Species: Mus musculus (house mouse)
C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S20303
R/Sundell, J.B.; Provvedini, D.M.; Wahlestedt, C.R.; Laurell, H.; Pohl, J.S.; Peterson, Eur. J. Biochem. 203, 625-631, 1992
A/Title: Molecular cloning of the murine substance K and substance P receptor genes.
A/Reference number: S20303; MUID:92137253; PMID:1170937
A/Accession: S20303
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-384 <SUN>
A/Cross-references: UNIPROT:P30549; UNIPARC:UPI0000029851; GB:X62933; NID:G54204; PIDN:C
C/Superfamily: neurokinin 1 receptor
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 68.4%; Score 78; DB 2; Length 384;
Best Local Similarity 60.9%; Pred. No. 6.3e-05;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 NYPLVNLAFASMAAFNTVNF 23
|||:|||||:|||||
Db 69 NYPIINLALADLCMAAFNATFNF 91

RESULT 14

A36737
neurokinin 2 receptor - rat
N/Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004
C/Accession: A36737
R/Saebel, Y.; Nakanishi, S. Biochem. Biophys. Res. Commun. 165, 695-702, 1989
A/Title: Molecular characterization of rat substance K receptor and its mRNA.
A/Reference number: A36737; MUID:90088481; PMID:2480781
A/Accession: A36737
A/Molecule type: mRNA
A/Residues: 1-390 <SAS>
A/Cross-references: UNIPROT:P16610; UNIPARC:UPI00001301F3; GB:M31838; NID:G206986; PIDN:C
C/Superfamily: neurokinin 1 receptor
C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 68.4%; Score 78; DB 2; Length 390;
Best Local Similarity 60.9%; Pred. No. 6.4e-05;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 NYPLVNLAFASMAAFNTVNF 23
|||:|||||:|||||
Db 69 NYPIINLALADLCMAAFNATFNF 91

RESULT 15

T32517
hypothetical protein C49A9.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T32517
R/Fulton, B.; Wohldmann, P. submitted to the EMBL Data Library, December 1997
A/Description: The sequence of C. elegans coemid C49A9.
A/Reference number: Z21184
A/Accession: T32517
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-391 <FUL>

A/Cross-references: UNIPROT:O44148; UNIPARC:UPI000003C147; EMBL:AF036693; PIDN:AAB88331.
A/Experimental source: strain Bristol N2; clone C49A9
C/Genetics:
A/Gene: C49A9.7
A/Map position: 4
A/Introns: 21/3; 84/2; 128/2; 159/2; 159/3; 233/2; 263/3; 347/2
C/Superfamily: neurokinin 1 receptor

Query Match 57.0%; Score 65; DB 2; Length 391;
Best Local Similarity 63.2%; Pred. No. 0.0085;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYPLVNLAFASMAAFNT 19
|||:|||||:|||||
Db 97 NYPLVNLAVADASIVENT 115

Search completed: January 23, 2006, 10:43:09
Job time: 13.5546 secs

Page Blank (ustof)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:31:19 ; Search time 73.3188 Seconds
(without alignments)
221.323 Million cell updates/sec

Title: US-10-501-838a-24
Perfect score: 114
Sequence: 1 NYFLVNLAFASMAAFNTVNF 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	195	2	Q53TQ2_HUMAN
2	114	100.0	311	2	Q8RIK6_CAVPO
3	114	100.0	406	2	Q4VBI8_HUMAN
4	114	100.0	407	1	NKIR_CANPA
5	114	100.0	407	1	NKIR_CAVPO
6	114	100.0	407	1	NKIR_HUMAN
7	114	100.0	407	1	NKIR_MERUN
8	114	100.0	407	2	Q547C3_CAVPO
9	111	97.4	309	2	Q9DCJ9_BUFWA
10	111	97.4	371	2	O800X0_BUFWA
11	111	97.4	390	2	O8UFP8_BUFWA
12	111	97.4	411	2	Q9WEI3_CHICK
13	110	96.5	333	2	Q7T078_FUGRU
14	110	96.5	406	2	Q4SIU0_TETNG
15	110	96.5	408	1	NKIR_PANCA
16	109	95.6	407	1	NKIR_MOUSE
17	109	95.6	407	1	NKIR_RAT
18	109	95.6	407	2	Q8BYR7_MOUSE
19	107	93.9	439	2	Q4RGJ3_TETNG
20	105	92.1	464	2	Q4VBI9_HUMAN
21	105	92.1	465	1	NKIR_HUMAN
22	105	92.1	467	1	NKIR_RABIT
23	101	88.6	452	1	NKIR_MOUSE
24	100	87.7	452	1	NKIR_RAT
25	99	86.8	360	2	Q4SDK8_TETNG
26	97	85.1	452	2	Q6NXX1_MOUSE
27	93	81.6	372	2	Q4RJB2_TETNG
28	91	79.8	349	2	Q4SI00_TETNG
29	85	74.6	157	2	Q9ZSR5_CAVPO
30	85	74.6	140	1	NKIR_CAVPO
31	79	69.3	130	2	Q96KE0_HUMAN

32	79	69.3	384	1	NK2R_BOVIN	P05563 bos taurus
33	79	69.3	384	1	NK2R_CANPA	Q5ubh2 canis fam1
34	79	69.3	398	1	NK2R_HUMAN	P21452 homo sapien
35	79	69.3	398	2	Q8NGQ8_HUMAN	Q8ngq8 homo sapien
36	79	69.3	398	2	Q9UDE6_HUMAN	Q9ude6 homo sapien
37	79	69.3	398	2	Q4ORIS_HUMAN	Q4or15 homo sapien
38	79	69.3	384	1	NK2R_CAVPO	Q64077 cavia porce
39	78	68.4	384	1	NK2R_MESAU	P51144 mesocricetu
40	78	68.4	384	1	NK2R_MOUSE	P30549 mus musculu
41	78	68.4	384	2	Q5DUB0_MERUN	Q5dub0 meriones un
42	78	68.4	384	2	Q8BZV9_MOUSE	Q8bzv9 mus musculu
43	78	68.4	390	1	NK2R_RAT	P16610 rattus norv
44	78	68.4	431	1	Q8RBD1_UREUN	Q8rbd1 urechis uni
45	75	65.8	384	1	NK2R_RABIT	P79218 oryctolagus

ALIGNMENTS

RESULT 1
Q53TQ2_HUMAN PRELIMINARY; PRT; 195 AA.
ID Q53TQ2; AC Q53TQ2; DT 13-SEP-2005 (TREMblrel. 31, Created)
DT 13-SEP-2005 (TREMblrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMblrel. 31, Last annotation update)
DE Hypothetical protein TACR1 (Fragment).
GN Name=TACR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Rohlfing T., Maupin R., Hawkins M., Davidson T.;
RT The sequence of Homo sapiens BAC clone RP11-355F16."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DR EMBL: AC007681; AAY14950.1; -; Genomic DNA.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005886; C: plasma membrane; IEA.
DR GO: GO:0005184; F: rhodopsin-like receptor activity; IEA.
DR GO: GO:0004995; F: tachykinin receptor activity; IEA.
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . . IEA.
DR InterPro: IPR000276; GPCR Rhodopsin.
DR InterPro: IPR001681; Neurokinin receptor.
DR InterPro: IPR000046; NK1 receptor.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHOPOPSN.
DR PRINTS: PR01024; NEUROKININR.
DR PRINTS: PR00244; NEUROKININR.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Hypothetical protein; Receptor;
KW Transducer; Transmembrane.
FT NON TER 195
SQ SEQUENCE 195 AA; 22162 MW; B67479B08AE41ECB CRC64;
Query Match 100.0%; Score 114; DB 2; Length 195;

Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFABASMAAFTVNF 23
DB 68 NYFLVNLAFABASMAAFTVNF 90

RESULT 2

Q8R1R6_CAVPO PRELIMINARY; PRT; 311 AA.
AC Q8R1R6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 24, Last annotation update)
DE Truncated neurokinin-1 receptor.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Hystricognathi; Caviidae; Cavia.
CC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Hystricognathi; Caviidae; Cavia.
NCBI_TaxID=10141;

NUCLEOTIDE SEQUENCE.

RP STRAIN-Hartley IMVS; TISSUE=Coeliac ganglion;
RX MEDLINE=22540921; PubMed=12654513; DOI=10.1016/S0169-328X(03)00002-0;
RA Baker S.J., Morris J.L., Gibbins I.L.,
RT "Cloning of a C-terminally truncated NK-1 receptor from guinea-pig nervous system";
RT Brain Res. Mol. Brain Res. 111:136-147(2003).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
EMBL: AF428266; AB096431; -. mRNA.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0005886; C:Plasma membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0004955; F:tachykinin receptor activity; IEA.
DR GO: GO:0007165; P:G-protein coupled receptor protein signaln. . .; IEA.
DR GO: GO:0007276; GPCR Rhodopsin.
DR InterPro: IPR000276; GPCR Rhodopsin.
DR InterPro: IPR001681; Neurokinin receptor.
DR InterPro: IPR000046; NK1_receptor.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHOPOSPN.
DR PRINTS: PR01024; NEUROKININR.
DR PRINTS: PS00244; NEUROKININR.
DR PROSITE: PS00237; G_PROTEIN_RECCEP_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECCEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 311 AA; 35702 MW; 06DD67BCBFA0239 CRC64;

Query Match 100.0%; Score 114; DB 2; Length 311;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFABASMAAFTVNF 23
DB 68 NYFLVNLAFABASMAAFTVNF 90

RESULT 3

Q4VBL8_HUMAN PRELIMINARY; PRT; 406 AA.
AC Q4VBL8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE TAC1 protein (Fragment).
GN Name=TAC1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9615;

OX NCBI_TaxID=9606;

NUCLEOTIDE SEQUENCE.

RP TISSUE=G-protein coupled receptors;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussid T.B., Toshiyuki S., Abramson R.D., Mullaly S.J.,
RA Raha S.S., Loquellano N.A., Peters G.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gutarra P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

NUCLEOTIDE SEQUENCE.

RP TISSUE=G-protein coupled receptors;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
EMBL: BC055528; AA055528.1; -. mRNA.
DR InterPro: IPR00276; GPCR Rhodopsin.
DR InterPro: IPR001681; Neurokinin receptor.
DR InterPro: IPR000046; NK1_receptor.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHOPOSPN.
DR PRINTS: PR01024; NEUROKININR.
DR PRINTS: PS00244; NEUROKININR.
DR PROSITE: PS00237; G_PROTEIN_RECCEP_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECCEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 406 AA; 46036 MW; 65FEB8AE4DA7A1CC CRC64;

Query Match 100.0%; Score 114; DB 2; Length 406;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFABASMAAFTVNF 23
DB 67 NYFLVNLAFABASMAAFTVNF 89

RESULT 4

NC1R_CANFA STANDARD; PRT; 407 AA.
AC Q5DUB3;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Substance-P receptor (SPR) (NK-1 receptor) (NK-1R) (Tachykinin receptor 1).
GN Name=TACR1;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Placentalia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]

NUCLEOTIDE SEQUENCE [mRNA].

RA Engherg S., Dimota T.;
 RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
 CC substance P. It is probably associated with G proteins that
 CC activate a phosphatidylinositol-calcium second messenger system
 CC (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL, AJ84915; C158656.1; -; mRNA.
 CC InterPro; IPR000276; GPCR_Rhodopsin.
 CC InterPro; IPR001681; Neurokin_receptor.
 CC InterPro; IPR000046; NK1_receptor.
 CC Pfam; PF00001; 7tm.1; 1.
 CC PRINTS; PR00237; GPCRHHODOPSN.
 CC PRINTS; PR01024; NEUROKININR.
 CC PRINTS; PR00244; NEUROKININR.
 CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
 CC Phosphorylation; Receptor; Transducer; Transmembrane.
 CC TOPO_DOM 1 31
 CC TRANSMEM 32 54
 CC TOPO_DOM 55 64
 CC TRANSMEM 65 86
 CC TOPO_DOM 87 106
 CC TRANSMEM 107 128
 CC TOPO_DOM 129 148
 CC TRANSMEM 149 169
 CC TOPO_DOM 170 194
 CC TRANSMEM 195 219
 CC TOPO_DOM 220 248
 CC TRANSMEM 249 270
 CC TOPO_DOM 271 283
 CC TRANSMEM 284 308
 CC TOPO_DOM 309 407
 CC LIPID 322 322
 CC CARBOHYD 14 14
 CC DISULFID 18 18
 CC FT 105 180
 CC SQ SEQUENCE 407 AA; 46464 MW; 18CBF4BAFB4F14 CRC64;
 CC
 CC Query Match 100.0%; Score 114; DB 1; Length 407;
 CC Best Local Similarity 100.0%; Pred. No. 2.8e-09;
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 NYFLVNLAFAPASMAAENTVNF 23
 CC 68 NYFLVNLAFAPASMAAENTVNF 90
 CC
 CC RESULT 5
 CC ID NK1R_CAVPO STANDARD; PRT; 407 AA.
 CC AC P30547;
 CC DT 01-APR-1993 (Rel. 25, Created)
 CC DT 01-APR-1993 (Rel. 25, Last sequence update)
 CC DT 13-SEP-2005 (Rel. 48, Last annotation update)
 CC DE Substance-P receptor (SPR) (NK-1 receptor) (NK-1R) (Tachykinin
 CC receptor 1).
 CC GN Name=TACR1; Synonym=TAC1R;
 CC OS Cavia porcellus (Guinea pig).
 CC OC Bkaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 CC OC Hystricognathi; Caviidae; Cavia.
 CC OC NCBI_TaxID=10141;
 CC RN [1]

RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=uterus;
 RA MEDLINE=92256498; PubMed=1374648; DOI=10.1016/0167-4781(92)90105-9;
 RX Gorbulay V., Akhundova A., Luzius H., Fahrholz F.;
 RT "Molecular cloning of substance P receptor cDNA from guinea-pig
 RT uterus".
 RL Biochim. Biophys. Acta 1131:99-102 (1992).
 CC -1- FUNCTION: This is a receptor for the tachykinin neuropeptide
 CC substance P. It is probably associated with G proteins that
 CC activate a phosphatidylinositol-calcium second messenger system.
 CC The rank order of affinity of this receptor to tachykinins is:
 CC substance P > substance K > neuromedin K.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL, X64323; CAA45608.1; -; mRNA.
 CC PIR; S23510; S23510.
 CC InterPro; IPR000276; GPCR_Rhodopsin.
 CC InterPro; IPR001681; Neurokin_receptor.
 CC InterPro; IPR000046; NK1_receptor.
 CC Pfam; PF00001; 7tm.1; 1.
 CC PRINTS; PR00237; GPCRHHODOPSN.
 CC PRINTS; PR01024; NEUROKININR.
 CC PRINTS; PR00244; NEUROKININR.
 CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
 CC Phosphorylation; Receptor; Transducer; Transmembrane.
 CC TOPO_DOM 1 31
 CC TRANSMEM 32 54
 CC TOPO_DOM 55 64
 CC TRANSMEM 65 86
 CC TOPO_DOM 87 106
 CC TRANSMEM 107 128
 CC TOPO_DOM 129 148
 CC TRANSMEM 149 169
 CC TOPO_DOM 170 194
 CC TRANSMEM 195 219
 CC TOPO_DOM 220 248
 CC TRANSMEM 249 270
 CC TOPO_DOM 271 283
 CC TRANSMEM 284 308
 CC TOPO_DOM 309 407
 CC LIPID 322 322
 CC CARBOHYD 14 14
 CC DISULFID 18 18
 CC FT 105 180
 CC SQ SEQUENCE 407 AA; 46257 MW; F8CSD31C4BB4556B CRC64;
 CC
 CC Query Match 100.0%; Score 114; DB 1; Length 407;
 CC Best Local Similarity 100.0%; Pred. No. 2.8e-09;
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 NYFLVNLAFAPASMAAENTVNF 23
 CC 68 NYFLVNLAFAPASMAAENTVNF 90
 CC
 CC RESULT 6
 CC ID NK1R_HUMAN STANDARD; PRT; 407 AA.
 CC AC P25103;
 CC DT 01-MAY-1992 (Rel. 22, Created)
 CC DT 01-MAY-1992 (Rel. 22, Last sequence update)
 CC DT 13-SEP-2005 (Rel. 48, Last annotation update)
 CC DE Substance-P receptor (SPR) (NK-1 receptor) (NK-1R) (Tachykinin
 CC receptor 1).
 CC GN Name=TACR1; Synonym=TAC1R;
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 CC OC Hystricognathi; Caviidae; Cavia.
 CC OC NCBI_TaxID=9606;
 CC RN [1]

GN Name=TACR1; Synonyms=NK1R, TAC1R;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92028856; PubMed=1718267;
 RA Takeda Y., Chou K.B., Takeda J., Sachais B.S., Krause J.E.;
 RT "Molecular cloning, structural characterization and functional
 expression of the human substance P receptor.";
 RL Biochem. Biophys. Res. Commun. 179:1232-1240(1991).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RC TISSUE=Lung;
 RX MEDLINE=92062052; PubMed=1659396;
 RA Hopfner B., Powell S.J., Danke P., Briggs I., Graham A.;
 RT "Isolation and characterization of the human lung NK-1 receptor
 cDNA.";
 RL Biochem. Biophys. Res. Commun. 180:1110-1117(1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Placenta;
 RX MEDLINE=92201186; PubMed=1312928;
 RA Takehashi K., Tanaka A., Hara M., Nakaneishi S.;
 RT "The primary structure and gene organization of human substance P and
 neuropeptide K receptors.";
 RL Eur. J. Biochem. 204:1025-1033(1992).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92031510; PubMed=1657150;
 RA Gerard N.P., Garraway L.A., Eddy R.L. Jr., Shows T.B., Iijima H.,
 RA Paquet J.L., Gerard C.;
 RT "Human substance P receptor (NK-1): organization of the gene,
 chromosome localization, and functional expression of cDNA clones.";
 RL Biochemistry 30:10640-10646(1991).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92123148; PubMed=1310144;
 RA Fong T.M., Anderson S.A., Yu H., Huang R.-R.C., Strader C.D.;
 RT "Differential activation of intracellular effector by two isoforms of
 human neurokinin-1 receptor.";
 RL Mol. Pharmacol. 41:24-30(1992).
 RN [6]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Brain;
 RA Kopatz S.A., Aronstam R.S., Sharma S.V.;
 RT "cDNA clones of human proteins involved in signal transduction
 sequenced by the Guthrie cDNA resource center (www.cdna.org).";
 RL Submitted (NOV-2003) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheef C.F., Bhat N.K.,
 RA Hopfner B.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughran J.N., Peters G.J., Abramson R.D., Mullen J.S.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pany J., Helton B., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.B.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SHOWS THAT THERE IS NO SHORT SPLICED ISOFORM.
 RX PubMed=11864635; DOI=10.1016/S0014-2999(02)01278-5;
 RA Page N.M., Bell N.U.;
 RT "The human tachykinin NK1 (short form) and tachykinin NK4 receptor: a
 reanalysis.";
 RL Eur. J. Pharmacol. 437:27-30(2002).
 RN [9]
 RP BINDING TO ANTAGONIST CP 96345.
 RX MEDLINE=93205121; PubMed=8384323; DOI=10.1038/362350a0;
 RA Fong T.M., Cascieri M.A., Yu H., Bansal A., Swain C., Strader C.D.;
 RT "Amino-aromatic interaction between histidine 197 of the neurokinin-1
 receptor and CP 96345.";
 RL Nature 362:350-353(1993).
 CC -I- FUNCTION: This is a receptor for the tachykinin neuropeptide
 substance P. It is probably associated with G proteins that
 activate a phosphatidylinositol-calcium second messenger system.
 CC The rank order of affinity of this receptor to tachykinins is:
 CC substance P > substance K > neuropeptide K.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; S62045; AAB20168.2; -; mRNA.
 DR EMBL; M74290; AAA60601.1; -; mRNA.
 DR EMBL; M81797; AAA59933.1; -; mRNA.
 DR EMBL; X65177; CAA46292.1; -; Genomic DNA.
 DR EMBL; X65178; CAA46292.1; -; JOINED; Genomic DNA.
 DR EMBL; X65179; CAA46292.1; JOINED; Genomic DNA.
 DR EMBL; X65180; CAA46292.1; JOINED; Genomic DNA.
 DR EMBL; X65181; CAA46292.1; JOINED; Genomic DNA.
 DR EMBL; M76675; AAA59936.1; -; mRNA.
 DR EMBL; M84425; AAA36641.1; -; mRNA.
 DR EMBL; M84426; AAA36644.1; -; mRNA.
 DR EMBL; AY462098; AAR23925.1; -; mRNA.
 DR EMBL; BC074911; AAH74911.1; -; mRNA.
 DR EMBL; BC074912; AAH74912.1; -; mRNA.
 DR PIR; A41134; J01274.
 DR Ensembl; ENSG00000115353; Homo sapiens.
 DR HGN; HGNC:11526; TACR1.
 DR MIM; 162323;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0004995; F:tachykinin receptor activity; TAS.
 DR GO; GO:0009582; P:detection of abiotic stimulus; TAS.
 DR GO; GO:0007200; P:G-protein signaling, coupled to IP3 second . . . ; TAS.
 DR GO; GO:0006954; P:inflammatory response; TAS.
 DR GO; GO:0007638; P:mechanosensory behavior; TAS.
 DR GO; GO:0007217; P:tachykinin signaling pathway; TAS.
 DR InterPro; IPR000276; GPCR Rhodopsin.
 DR InterPro; IPR001681; Neurokinin receptor.
 DR InterPro; IPR000046; NK1_receptor.
 DR Pfam; PF00001; 7tm_1.1.
 DR PRINTS; PR00237; GPCR_RHODOPSN.
 DR PRINTS; PR01024; NEUROKININR.
 DR PRINTS; PR00244; NEUROKININR.
 DR PROSITE; PS00237; G_PROTEIN_RECPT_F1_1; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECPT_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate;
 KW Phosphorylation; Receptor; Transducer; Transmembrane.
 FT TOPO_DOM 1 31
 FT TRANSMEM 32 54
 FT TOPO_DOM 55 64
 FT TRANSMEM 65 86
 FT TOPO_DOM 87 106
 FT TRANSMEM 107 128
 FT TOPO_DOM 129 148
 FT Cytoplasmic (Potential).
 FT Extracellular (Potential).
 FT 3 (Potential).
 FT Cytoplasmic (Potential).

```

FT TRANSMEM 149 169 4 (Potential).
FT TOPO_DOM 170 194 Extracellular (Potential).
FT TRANSMEM 195 219 5 (Potential).
FT TOPO_DOM 220 248 Cytoplasmic (Potential).
FT TRANSMEM 249 270 6 (Potential).
FT TOPO_DOM 271 283 Extracellular (Potential).
FT TRANSMEM 284 308 7 (Potential).
FT TOPO_DOM 309 407 Cytoplasmic (Potential).
FT BINDING 197 197 Antagonist CP 96345.
FT LIPID 322 332 S-palmitoyl cysteine (Potential).
FT CARBOHYD 14 14 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 18 18 N-linked (GlcNAc...) (Potential).
FT DISULFID 105 180 By similarity.
FT CONFLICT 116 116 V -> C (in Ref. 4).
FT CONFLICT 218 218 V -> I (in Ref. 4).
SQ SEQUENCE 407 AA; 46251 MW; 2AFPD3F61BA3041 CRC64;

Query Match 100.0%; Score 114; DB 1; Length 407;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFARASMAFNTVNP 23
DB 68 NYFLVNLAFARASMAFNTVNP 90

RESULT 7
NK1R_MERUN STANDARD; PRT; 407 AA.
ID NK1R_MERUN
AC Q5DUB1;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DB Substance-P receptor (SPR) (NK-1 receptor) (NK-1R) (Tachykinin receptor 1).
GN Name-TACR1;
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Gerbillinae; Meriones.
OC NCBI_TaxId=10047;
OX NCBI_TaxId=10047;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RA Engberg S., Dimota T.;
RL Submitted (FE8-2005) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This is a receptor for the tachykinin neuropeptide substance P. It is probably associated with G proteins that activate a phosphatidylinositol-calcium second messenger system (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC EMBL; AUB84917; CA158658.1; -; mRNA.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR000461; NK1_receptor.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PRINTS; PR01024; NEUROKININR.
DR PRINTS; PR00237; G_PROTEIN_RECCEP_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECCEP_FL_2; 1.
KW G-protein coupled receptor; Glycoprotein; Lipoprotein; Palmitate; Phosphorylation; Receptor; Transducer; Transmembrane.
FT TOPO_DOM 1 31 Extracellular (Potential).
FT TRANSMEM 32 54 1 (Potential).
FT TOPO_DOM 55 64 Cytoplasmic (Potential).

```

```

FT TRANSMEM 65 86 2 (Potential).
FT TOPO_DOM 87 106 Extracellular (Potential).
FT TRANSMEM 107 128 3 (Potential).
FT TOPO_DOM 129 148 Cytoplasmic (Potential).
FT TRANSMEM 149 169 4 (Potential).
FT TOPO_DOM 170 194 Extracellular (Potential).
FT TRANSMEM 195 219 5 (Potential).
FT TOPO_DOM 220 248 Cytoplasmic (Potential).
FT TRANSMEM 249 270 6 (Potential).
FT TOPO_DOM 271 283 Extracellular (Potential).
FT TRANSMEM 284 308 7 (Potential).
FT TOPO_DOM 309 407 Cytoplasmic (Potential).
FT LIPID 322 332 S-palmitoyl cysteine (Potential).
FT CARBOHYD 14 14 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 18 18 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT DISULFID 189 189 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 407 AA; 46178 MW; 1AB2ED9632325023 CRC64;

Query Match 100.0%; Score 114; DB 1; Length 407;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFARASMAFNTVNP 23
DB 68 NYFLVNLAFARASMAFNTVNP 90

RESULT 8
Q547C3_CAVPO PRELIMINARY; PRT; 407 AA.
ID Q547C3;
AC Q547C3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DB Neurokinin-1 receptor.
GN Neurokinin-1 receptor.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Caviidae; Cavia.
OC NCBI_TaxId=10141;
OX NCBI_TaxId=10141;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA TISSUE-Coeliac ganglion;
RX MEDLINE=22540921; PubMed=12654513; DOI=10.1016/S0169-328X(03)00002-0; Baker S.J., Morris J.L., Gibbins J.L.;
RA "Cloning of a C-terminally truncated NK-1 receptor from guinea-pig nervous system.";
RL Brain Res. Mol. Brain Res. 111:136-147(2003).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC EMBL; AF509591; AM34738.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005886; C:plasma membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0004584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0004995; F:tachykinin receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR001681; Neurokinin_receptor.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PRINTS; PR01024; NEUROKININR.
DR PRINTS; PR00237; G_PROTEIN_RECCEP_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECCEP_FL_2; 1.
KW G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 407 AA; 46257 MW; F8C5D31C4B4586B CRC64;

Query Match 100.0%; Score 114; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;

```



```

RESULT 12
Q9W613 CHICK PRELIMINARY, PRT, 411 AA.
AC Q9W613_
DT 01-NOV-1999 (TRENBLREL. 12, Created)
DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Substance P receptor.
GN Name=ASPR;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISUB=Brain;
RA Sia G.M., Magglo J.B., Too H.P.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DR EMBL, AF131057; AAD31017.1; -; mRNA. Gallus.
DR Ensembl; ENSGALG00000013953; Gallus gallus.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005886; C:plasma membrane; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO; GO:0004995; P:tyrosine kinase receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR001681; Neurokinin receptor.
DR InterPro; IPR000046; NK1_receptor.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PRINTS; PR01024; NEUROKININR.
DR PRINTS; PR00244; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_REC_P1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_REC_P1_2; 1.
DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.
KW SEQUENCE 411 AA; 47091 MW; 35289DB6A046159 CRC64;
SQ

Query Match 97.4%; Score 111; DB 2; Length 411;
Best Local Similarity 95.7%; Pred. No. 8.4e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPYLVNLAFAEASMAAFNTVNF 23
DB 72 NPYLVNLAFAEASMAAFNTVNF 94

RESULT 13
Q70708 FUGRU PRELIMINARY, PRT, 393 AA.
AC Q70708_
DT 01-OCT-2003 (TRENBLREL. 25, Created)
DT 01-OCT-2003 (TRENBLREL. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
DE Putative tachykinin receptor 1.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OC NCBI_TaxID=31033;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Lagerstrom M.C., Klovins J., Fredriksson R., Fridmanis D., Haltina T.,
RA Ling M.K., Berling M.M., Schloch H.B.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

```

```

DR EMBL; AY327862; AA002694.1; -; Genomic DNA.
DR Ensembl; SINEUG00000151158; Fugu rubripes.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005886; C:plasma membrane; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO; GO:0004995; P:tyrosine kinase receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR00276; GPCR_Rhodopsin.
DR InterPro; IPR001681; Neurokinin receptor.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PRINTS; PR01024; NEUROKININR.
DR PRINTS; PR00244; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_REC_P1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_REC_P1_2; 1.
DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.
KW SEQUENCE 393 AA; 45115 MW; 7F508031ADF57F8 CRC64;
SQ

Query Match 96.5%; Score 110; DB 2; Length 393;
Best Local Similarity 91.3%; Pred. No. 1.2e-08;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NPYLVNLAFAEASMAAFNTVNF 23
DB 72 NPYLVNLAFAEASMAAFNTVNF 94

RESULT 14
Q4S1U0 TERN PRELIMINARY, PRT, 406 AA.
AC Q4S1U0_
DT 13-SEP-2005 (TRENBLREL. 31, Created)
DT 13-SEP-2005 (TRENBLREL. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLREL. 31, Last annotation update)
DE Chromosome undetermined SCAP14764, whole genome shotgun sequence.
GN ORFNames=GSTENG00025389001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OC NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Jallion O., Aubry J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Maucell E., Bouneau L., Fischer C., Orouf-Coetz C., Bernot A.,
RA Nicoud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skallil Z., Catolico L., Poulain J., De Berardinis V.,
RA Crnaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.R., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lauder V., Schacher V., Queller F., Sautin W., Scarpelli H.,
RA Wincker P., Lander B.S., Weissbach J., Roest Croli H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
DR EMBL; CAAB01014764; CG05392.1; -; Genomic DNA.
DR InterPro; IPR000276; GPCR_Rhodopsin.

```

```
DR InterPro: IPR001681; Neurokin receptor.
DR InterPro: IPR000046; NK1_receptor.
DR Pfam: PF00001; Tcm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01024; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Receptor; Transducer; Transmembrane.
SQ SEQUENCE 406 AA; 46319 MW; 40057E1BEA9157B9 CRC64;

Query Match
Best Local Similarity 96.5%; Score 110; DB 2; Length 406;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFASMSAFTVTVNF 23
|||||
72 NYFLVNLAFASMSAFTVTVNF 94

RESULT 15
ID NK1R_RANCA STANDARD; PRT; 408 AA.
AC Q98982;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Substance-P receptor (SPR) (NK-1 receptor) (NK-1R) (Tachykinin
DE receptor 1).
GN Name=TACR1; Synonyms=TAC1R;
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana;
OC Aquarana.
OX NCBI_TaxID=8400;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Sympathetic ganglion;
RX MEDLINE=97363687; PubMed=9219980; DOI=10.1016/S0306-4522(97)00027-4;
RA Simmons M.A., Brodbeck R.M., Karplitskiy V.V., Schneider C.R.,
RA Neff D.P.A., Krause J.B.;
RT "Molecular characterization and functional expression of a substance P
RT receptor from the sympathetic ganglion of Rana catesbeiana.";
RL Neuroscience 79:1219-1229(1997).
CC -!- FUNCTION: This is a receptor for the tachykinin neuropeptide
CC substance P. It is probably associated with G proteins that
CC activate a phosphatidylinositol-calcium second messenger system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; U67736; AAC05707.1; -; mRNA.
DR HSSP; P02699; IP88.
DR InterPro; IPR000276; GPCR_Rhodopsin.
DR InterPro; IPR001681; Neurokin receptor.
DR InterPro; IPR000046; NK1_receptor.
DR Pfam; PF00001; Tcm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01024; NEUROKININR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Glycophorin; Lipoprotein; Palmitate;
KW Phosphorylation; Receptor; Transducer; Transmembrane.
FT TOPO_DOM 1 32 Extracellular (Potential).
FT TRANSMEM 33 55 1 (Potential).
FT TOPO_DOM 56 65 Cytoplasmic (Potential).
FT TRANSMEM 66 87 2 (Potential).
```

```
FT TOPO_DOM 88 107 Extracellular (Potential).
FT TRANSMEM 108 129 3 (Potential).
FT TOPO_DOM 130 149 Cytoplasmic (Potential).
FT TRANSMEM 150 170 4 (Potential).
FT TOPO_DOM 171 196 Extracellular (Potential).
FT TRANSMEM 197 221 5 (Potential).
FT TOPO_DOM 222 250 Cytoplasmic (Potential).
FT TRANSMEM 251 272 6 (Potential).
FT TOPO_DOM 273 283 Extracellular (Potential).
FT TRANSMEM 284 308 7 (Potential).
FT TOPO_DOM 309 408 Cytoplasmic (Potential).
FT LIPID 323 323 S-palmitoyl cysteine (Potential).
FT CARBOHYD 4 4 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 9 9 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 14 14 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 19 19 N-linked (GlcNAc...) (Potential).
FT DISULFID 106 181 By similarity.
SQ SEQUENCE 408 AA; 47011 MW; 300C28A8E8820640E CRC64;

Query Match
Best Local Similarity 96.5%; Score 110; DB 1; Length 408;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFASMSAFTVTVNF 23
|||||
69 NYFLVNLAFASMSAFTVTVNF 91

Search completed: January 23, 2006, 10:40:56
Job time : 74.3188 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using sw model

Run on: January 23, 2006, 09:35:50 | Search time 21.0917 Seconds
(without alignments)

90.156 Million cell updates/sec

Title: US-10-501-838A-24

Sequence: 1 NYFLVNLAFASMAAFNTVNF 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/prodata/1/1aa/5/COMB.pep.*
2: /cgn2_6/prodata/1/1aa/6/COMB.pep.*
3: /cgn2_6/prodata/1/1aa/R/COMB.pep.*
4: /cgn2_6/prodata/1/1aa/PCTUS/COMB.pep.*
5: /cgn2_6/prodata/1/1aa/R/COMB.pep.*
6: /cgn2_6/prodata/1/1aa/Backfill.pep.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	311	1	US-07-701-935-26 Sequence 26, Appl
2	114	100.0	407	1	US-08-117-965-26 Sequence 26, Appl
3	114	100.0	407	2	US-09-826-509-525 Sequence 525, App
4	114	100.0	407	4	PCT-US92-06532-3 Sequence 3, Appl
5	109	95.6	372	1	US-07-937-609-20 Sequence 20, Appl
6	109	95.6	372	2	US-08-029-170-20 Sequence 20, Appl
7	109	95.6	372	2	US-09-443-745-20 Sequence 20, Appl
8	109	95.6	407	4	PCT-US92-06532-7 Sequence 7, Appl
9	105	92.1	465	2	US-08-090-369-1 Sequence 1, Appl
10	105	92.1	465	2	US-09-482-971-1 Sequence 1, Appl
11	105	92.1	465	2	US-09-826-509-529 Sequence 529, App
12	104	91.2	407	1	US-08-390-000A-6 Sequence 6, Appl
13	100	87.7	336	4	US-08-118-270-50 Sequence 50, Appl
14	100	87.7	336	4	PCT-US93-08528-50 Sequence 50, Appl
15	100	87.7	411	1	US-07-937-609-21 Sequence 21, Appl
16	100	87.7	411	2	US-08-029-170-21 Sequence 21, Appl
17	100	87.7	411	2	US-09-443-745-21 Sequence 21, Appl
18	100	87.7	451	2	US-08-430-286A-10 Sequence 10, Appl
19	85	74.6	440	2	US-08-430-286A-9 Sequence 9, Appl
20	79	69.3	384	1	US-08-103-170-10 Sequence 10, Appl
21	79	69.3	387	1	US-08-196-989B-14 Sequence 14, Appl
22	79	69.3	387	1	US-08-760-936-14 Sequence 14, Appl
23	79	69.3	387	2	US-09-225-024-14 Sequence 14, Appl
24	79	69.3	398	2	US-09-826-509-527 Sequence 527, App
25	78	68.4	369	1	US-07-937-609-19 Sequence 19, Appl
26	78	68.4	369	2	US-08-029-170-19 Sequence 19, Appl
27	78	68.4	369	2	US-09-443-745-19 Sequence 19, Appl

ALIGNMENTS

28	65	57.0	391	2	US-09-721-870-14	Sequence 34, Appl
29	62	54.4	519	2	US-09-693-746-10	Sequence 10, Appl
30	62	54.4	522	2	US-09-693-746-12	Sequence 12, Appl
31	60	52.6	341	1	US-08-118-270-48	Sequence 48, Appl
32	59	51.8	341	4	PCT-US93-08528-48	Sequence 48, Appl
33	59	51.8	374	2	US-09-721-870-24	Sequence 24, Appl
34	51	44.7	501	1	US-08-722-001-14	Sequence 14, Appl
35	51	44.7	501	1	US-08-467-568-9	Sequence 9, Appl
36	51	44.7	501	1	US-09-030-582-9	Sequence 9, Appl
37	51	44.7	501	2	US-09-688-415-7	Sequence 7, Appl
38	51	44.7	572	1	US-08-334-698-2	Sequence 2, Appl
39	51	44.7	572	1	US-08-228-932-2	Sequence 2, Appl
40	51	44.7	572	1	US-08-468-939-2	Sequence 2, Appl
41	51	44.7	572	1	US-08-722-001-30	Sequence 30, Appl
42	51	44.7	572	1	US-08-406-855A-2	Sequence 2, Appl
43	51	44.7	572	1	US-08-722-190-2	Sequence 2, Appl
44	51	44.7	572	2	US-08-244-354-2	Sequence 2, Appl
45	51	44.7	572	2	US-09-206-899-2	Sequence 2, Appl

RESULT 1
US-07-701-935-26
Sequence 26, Application US/07701935
Patent No. 5336595
GENERAL INFORMATION:
APPLICANT: Strader, C. D.
TITLE OF INVENTION: Method of Using Human Neurokinin-1
NUMBER OF INVENTIONS: 27
TITLE OF INVENTIONS: Receptor Short Form
CORRESPONDENCE ADDRESS:
ADDRESSER: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07701, 935
FILING DATE: 19910517
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nicholson, William H.
REGISTRATION NUMBER: 25,147
REFERENCE/DOCKET NUMBER: 18409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-5315
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 311 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
US-07-701-935-26
Query Match 100.0%; Score 114; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 1e-09; 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;
CY 1 NYFLVNLAFASMAAFNTVNF 23
DB 68 NYFLVNLAFASMAAFNTVNF 90

```
RESULT 2
US-08-117-965-26
; Sequence 26, Application US/08117965
; Patent No. 5484886
; GENERAL INFORMATION:
; APPLICANT: Tung, Fong M.
; TITLE OF INVENTION: Human Neurokinin-1 Receptor
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P. O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,965
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 691,197
; FILING DATE: 25-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Nicholson, William H.
; REGISTRATION NUMBER: 25,147
; REFERENCE/DOCKET NUMBER: 18387
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-5315
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 407 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; US-08-117-965-26

Query Match          100.0%; Score 114; DB 1; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NYFLVNLAFARASMAAFNTVNF 23
DB      68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 3
US-09-826-509-525
; Sequence 525, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; TITLE OF INVENTION: Protein-Coupled Receptors
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: Patentin Version 2.1
```

```
; SEQ ID NO 525
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-525

Query Match          100.0%; Score 114; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NYFLVNLAFARASMAAFNTVNF 23
DB      68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 4
PCT-US92-06532-3
; Sequence 3, Application PC/TUS9206532
; GENERAL INFORMATION:
; APPLICANT: Krause, James B.
; TITLE OF INVENTION: Human Substance P Receptor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Scott J. Meyer, Monsanto Co., ASD
; STREET: 800 N. Lindbergh Blvd.
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: U.S.A
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06532
; FILING DATE: 19920805
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: 07-24 (776)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)694-3117
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 407 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-06532-3

Query Match          100.0%; Score 114; DB 4; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NYFLVNLAFARASMAAFNTVNF 23
DB      68 NYFLVNLAFARASMAAFNTVNF 90

RESULT 5
US-07-937-609-20
; Sequence 20, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
```

CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIH
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance P receptor
US-07-937-609-20

Query Match 95.6%; Score 109; DB 1; Length 372;
Best Local Similarity 95.7%; Pred. No. 7.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYFLVNLAFARASMAAFNTVNF 23
DB 68 NYFLVNLAFARACMAAFNTVNF 90

RESULT 6
US-08-029-170-20
Sequence 20, Application US/08029170
Patent No. 6168173
GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIH
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance P receptor
US-08-029-170-20

Query Match 95.6%; Score 109; DB 2; Length 372;
Best Local Similarity 95.7%; Pred. No. 7.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYFLVNLAFARASMAAFNTVNF 23
DB 68 NYFLVNLAFARACMAAFNTVNF 90

RESULT 7
US-09-443-745-20
Sequence 20, Application US/09443745
Patent No. 6706493
GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769

FILED DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NTHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance P receptor
US-09-443-745-20

Query Match 95.6%; Score 109; DB 2; Length 372;
Best Local Similarity 95.7%; Pred. No. 7.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFAPASMAAFNTVNF 23
Db 68 NYFLVNLAFAPACMAAFNTVNF 90

RESULT 8
PCT-US92-06532-7
Sequence 7, Application PC/TUS9206532
GENERAL INFORMATION:
APPLICANT: Krause, James E.
TITLE OF INVENTION: Human Substance P Receptor
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SD
STREET: 800 N. Lindbergh Blvd.
CITY: St. Louis
STATE: Missouri
COUNTRY: U.S.A
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06532
FILING DATE: 19920805
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: 07-24(776)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)694-3117
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 407 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-06532-7

Query Match 95.6%; Score 109; DB 4; Length 407;
Best Local Similarity 95.7%; Pred. No. 8e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFAPASMAAFNTVNF 23

Db 68 NYFLVNLAFAPACMAAFNTVNF 90

RESULT 9
US-08-090-369-1
Sequence 1, Application US/08090369
Patent No. 6258943
GENERAL INFORMATION:
APPLICANT: Fong, T.M.
APPLICANT: Huang, R-R. C.
APPLICANT: Strader, C.D.
TITLE OF INVENTION: Human Neurokinin-3 Receptor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: USA
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/090,369
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/851,974
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Thies, J. E.
REGISTRATION NUMBER: P-35,382
REFERENCE/DOCKET NUMBER: 18685
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3904
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-090-369-1

Query Match 92.1%; Score 105; DB 2; Length 465;
Best Local Similarity 87.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFAPASMAAFNTVNF 23
Db 121 NYFLVNLAFSDASMAAFNTLVNF 143

RESULT 10
US-09-482-971-1
Sequence 1, Application US/09482971
Patent No. 6348330
GENERAL INFORMATION:
APPLICANT: Fong, T.M.
APPLICANT: Huang, R-R. C.
APPLICANT: Strader, C.D.
TITLE OF INVENTION: Human Neurokinin-3 Receptor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey

COUNTRY: USA
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/482,971
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/090,369
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Thies, J. E.
REGISTRATION NUMBER: P-35,382
REFERENCE/DOCKET NUMBER: 18685
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3904
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-482-971-1

Query Match 92.1%; Score 105; DB 2; Length 465;
Best Local Similarity 87.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPLVNLAFPAASMAAFNTVNF 23
DB 121 NYPLVNLAFSPDASMAAFNTVNF 143

RESULT 11
US-09-826-509-529
Sequence 529, Application US/09826509
Patent No. 6806054
GENERAL INFORMATION:
APPLICANT: Lehmann-Brulnema, Karin
APPLICANT: Lfaw, Chen W.
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: Patentin Version 2.1
SEQ ID NO 529
LENGTH: 465
TYPE: PRT
ORGANISM: Homo sapiens
US-09-826-509-529

Query Match 92.1%; Score 105; DB 2; Length 465;
Best Local Similarity 87.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPLVNLAFPAASMAAFNTVNF 23
DB 121 NYPLVNLAFSPDASMAAFNTVNF 143

RESULT 12
US-08-390-000A-6
Sequence 6, Application US/08390000A
Patent No. 5985583
GENERAL INFORMATION:
APPLICANT: Sealton, Stuart C.
TITLE OF INVENTION: Cloning and Expression of
TITLE OF INVENTION: Gonadotropin-Releasing Hormone Receptor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,000A
FILING DATE: 17-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6923-052
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 407 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-390-000A-6

Query Match 91.2%; Score 104; DB 1; Length 407;
Best Local Similarity 91.3%; Pred. No. 4.7e-08;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYPLVNLAFPAASMAAFNTVNF 23
DB 68 NYPLVNLAFPAACMAAFNTVNF 90

RESULT 13
US-08-118-270-50
Sequence 50, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-50

Query Match 87.7%; Score 100; DB 1; Length 336;
Best Local Similarity 78.3%; Pred. No. 1.6e-07;
Matches 18; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPLVNLAFASMAAFNTVNF 23
DB 37 NYPLVNLAFSDASVAAPFTLINF 59

RESULT 14

PCT-US93-08528-50
Sequence 50, Application PC/TUS9308528

GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSER: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-50

Query Match 87.7%; Score 100; DB 4; Length 336;
Best Local Similarity 78.3%; Pred. No. 1.6e-07;
Matches 18; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPLVNLAFASMAAFNTVNF 23
DB 37 NYPLVNLAFSDASVAAPFTLINF 59

RESULT 15

US-07-937-609-21
Sequence 21, Application US/07937609
Patent No. 5319073

GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIH
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-937-609-21

Query Match 87.7%; Score 100; DB 1; Length 411;
Best Local Similarity 78.3%; Pred. No. 2e-07;
Matches 18; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPLVNLAFASMAAFNTVNF 23
DB 108 NYPLVNLAFSDASVAAPFTLINF 130

Search completed: January 23, 2006, 10:46:49
Job time : 21.0917 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 11:52:43 ; Search time 52.8297 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-24

Sequence: 1 NYFLVNLAFPAESMAAFNTVNF 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US11_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	100.0	23	4	US-10-665-184-24 Sequence 24, App1
2	114	100.0	23	5	US-10-942-300-24 Sequence 24, App1
3	114	100.0	23	5	US-10-501-838a-24 Sequence 24, App1
4	114	100.0	29	4	US-10-665-184-34 Sequence 34, App1
5	114	100.0	29	5	US-10-942-300-34 Sequence 34, App1
6	114	100.0	29	5	US-10-501-838a-34 Sequence 34, App1
7	114	100.0	311	3	US-09-930-503-1 Sequence 1, App1
8	114	100.0	311	3	US-09-930-503-3 Sequence 3, App1
9	114	100.0	407	3	US-09-930-503-5 Sequence 5, App1
10	114	100.0	407	3	US-09-930-503-7 Sequence 7, App1
11	114	100.0	407	3	US-09-826-509-525 Sequence 525, App
12	114	100.0	407	4	US-10-005-956-14 Sequence 14, App1
13	114	100.0	407	4	US-10-005-956-16 Sequence 16, App1
14	114	100.0	407	4	US-10-005-956-18 Sequence 18, App1
15	114	100.0	407	4	US-10-005-956-20 Sequence 20, App1
16	114	100.0	407	4	US-10-005-956-22 Sequence 22, App
17	114	100.0	407	4	US-10-225-567A-322 Sequence 322, App
18	114	100.0	407	5	US-10-925-095-525 Sequence 525, App
19	114	100.0	407	5	US-10-451-304-5 Sequence 5, App1
20	114	100.0	407	5	US-10-451-304-18 Sequence 18, App1
21	114	97.4	410	4	US-10-254-905-8 Sequence 8, App1
22	111	97.4	411	3	US-09-966-782A-8 Sequence 8, App1
23	109	95.6	372	3	US-09-443-745-20 Sequence 20, App1
24	105	92.1	465	4	US-09-826-509-529 Sequence 529, App
25	105	92.1	465	4	US-10-225-567A-197 Sequence 197, App
26	105	92.1	465	5	US-10-925-095-529 Sequence 529, App
27	105	92.1	488	4	US-10-029-009-9 Sequence 9, App1

28	105	92.1	505	4	US-10-029-009-21 Sequence 21, App1
29	100	87.7	411	3	US-09-443-745-21 Sequence 21, App1
30	85	74.6	436	4	US-10-254-905-7 Sequence 7, App1
31	85	74.6	440	3	US-09-966-782A-7 Sequence 7, App1
32	85	74.6	440	3	US-09-966-782A-7 Sequence 7, App1
33	83	72.8	440	3	US-10-225-567A-545 Sequence 545, App
34	79	69.3	138	4	US-10-029-386-32429 Sequence 32429, A
35	79	69.3	23	4	US-10-665-184-9 Sequence 9, App1
36	79	69.3	23	5	US-10-942-300-9 Sequence 9, App1
37	79	69.3	23	5	US-10-501-838a-9 Sequence 9, App1
38	79	69.3	30	4	US-10-665-184-31 Sequence 31, App1
39	79	69.3	30	5	US-10-942-300-31 Sequence 31, App1
40	79	69.3	313	4	US-10-501-838A-31 Sequence 31, App1
41	79	69.3	313	4	US-10-060-795B-7 Sequence 7, App1
42	79	69.3	398	3	US-09-826-509-537 Sequence 527, App
43	79	69.3	398	4	US-10-225-567A-456 Sequence 456, App
44	79	69.3	398	4	US-10-160-358-3 Sequence 3, App1
45	79	69.3	398	4	US-10-292-798-648 Sequence 648, App
					Sequence 527, App

ALIGNMENTS

RESULT 1
US-10-665-184-24
; Sequence 24, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-184-24

Query Match 100.0%; Score 114; DB 4; Length 23;
Best Local Similarity 100.0%; Pred No. 4,5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFPAESMAAFNTVNF 23
DB 1 NYFLVNLAFPAESMAAFNTVNF 23

RESULT 2
US-10-942-300-24
; Sequence 24, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 24
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-942-300-24

Query Match 100.0%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 4,5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFAEASMAAFNTVNF 23
Db 1 NYFLVNLAFAEASMAAFNTVNF 23

RESULT 3
US-10-501-838A-24
; Sequence 24, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT FILING DATE: 2004-07-19
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-501-838A-24

Query Match 100.0%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 4,5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFAEASMAAFNTVNF 23
Db 1 NYFLVNLAFAEASMAAFNTVNF 23

RESULT 4
US-10-665-184-34
; Sequence 34, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial

;; FEATURE:
;; OTHER INFORMATION: Penetrating peptide
;; FEATURE:
;; NAME/KEY: MOD_RES
;; LOCATION: (1)..(1)
;; OTHER INFORMATION: ACETYLATION
;; FEATURE:
;; NAME/KEY: PEPTIDE
;; LOCATION: (29)..(29)
;; OTHER INFORMATION: wherein Xaa is Lysine-NH2
US-10-665-184-34

Query Match 100.0%; Score 114; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 5,8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFAEASMAAFNTVNF 23
Db 1 NYFLVNLAFAEASMAAFNTVNF 23

RESULT 5
US-10-942-300-34
; Sequence 34, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Acylated Penetrating Peptide
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (29)..(29)
; OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated
US-10-942-300-34

Query Match 100.0%; Score 114; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 5,8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYFLVNLAFAEASMAAFNTVNF 23
Db 1 NYFLVNLAFAEASMAAFNTVNF 23

RESULT 6
US-10-501-838A-34
; Sequence 34, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL


```
; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic; penetrating peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (26)..(26)
; OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (29)..(29)
; OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (29)..(29)
; OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
; OTHER INFORMATION: peptide via the free amino groups of the lysine residue
US-10-501-838A-34
```

```
Query Match          100.0%; Score 114; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 5.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 NYFLVNLAPFAASMAAFNTVNF 23
Db 1 NYFLVNLAPFAASMAAFNTVNF 23
```

```
RESULT 7
US-09-930-503-1
; Sequence 1, Application US/09930503
; Publication No. US20030060438A1
; GENERAL INFORMATION:
; APPLICANT: HENRY, JAMES
; APPLICANT: CAHILL, CATHERINE
; APPLICANT: YASHPAL, KIRAN
; TITLE OF INVENTION: APPLICATION OF ANTI-SENSE OLIGONUCLEOTIDES AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; FILE REFERENCE: 39245-173913
; CURRENT APPLICATION NUMBER: US/09/930,503
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/226,086
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-503-1
```

```
Query Match          100.0%; Score 114; DB 3; Length 311;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 NYFLVNLAPFAASMAAFNTVNF 23
Db 68 NYFLVNLAPFAASMAAFNTVNF 90
```

RESULT 8

```
US-09-930-503-3
; Sequence 3, Application US/09930503
; Publication No. US20030060438A1
; GENERAL INFORMATION:
; APPLICANT: HENRY, JAMES
; APPLICANT: CAHILL, CATHERINE
; APPLICANT: YASHPAL, KIRAN
; TITLE OF INVENTION: APPLICATION OF ANTI-SENSE OLIGONUCLEOTIDES AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; FILE REFERENCE: 39245-173913
; CURRENT APPLICATION NUMBER: US/09/930,503
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/226,086
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-503-3
```

```
Query Match          100.0%; Score 114; DB 3; Length 311;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 NYFLVNLAPFAASMAAFNTVNF 23
Db 68 NYFLVNLAPFAASMAAFNTVNF 90
```

```
RESULT 9
US-09-930-503-5
; Sequence 5, Application US/09930503
; Publication No. US20030060438A1
; GENERAL INFORMATION:
; APPLICANT: HENRY, JAMES
; APPLICANT: CAHILL, CATHERINE
; APPLICANT: YASHPAL, KIRAN
; TITLE OF INVENTION: APPLICATION OF ANTI-SENSE OLIGONUCLEOTIDES AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
; FILE REFERENCE: 39245-173913
; CURRENT APPLICATION NUMBER: US/09/930,503
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/226,086
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-503-5
```

```
Query Match          100.0%; Score 114; DB 3; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 NYFLVNLAPFAASMAAFNTVNF 23
Db 68 NYFLVNLAPFAASMAAFNTVNF 90
```

```
RESULT 10
US-09-930-503-7
; Sequence 7, Application US/09930503
; Publication No. US20030060438A1
; GENERAL INFORMATION:
; APPLICANT: HENRY, JAMES
; APPLICANT: CAHILL, CATHERINE
; APPLICANT: YASHPAL, KIRAN
; TITLE OF INVENTION: APPLICATION OF ANTI-SENSE OLIGONUCLEOTIDES AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREOF
```

```
FILE REFERENCE: 39245-173913
; CURRENT APPLICATION NUMBER: US/09/930,503
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/226,086
; PRIOR FILING DATE: 2000-08-18
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 7
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-503-7

Query Match          100.0%; Score 114; DB 3; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFAEASMAAFNTVNF 23
    |||||
Db 68 NYFLVNLAFAEASMAAFNTVNF 90

RESULT 11
US-09-826-509-525
; Sequence 525, Application US/09826509
; Publication No. US20030204073A1
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Brunema, Karin
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. US20030204073A1-Endogenous, Constitutively Activated Known G
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO: 525
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-525

Query Match          100.0%; Score 114; DB 3; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFAEASMAAFNTVNF 23
    |||||
Db 68 NYFLVNLAFAEASMAAFNTVNF 90

RESULT 12
US-10-005-956-14
; Sequence 14, Application US/10005956
; Publication No. US20030113726A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: D0053NP
; CURRENT APPLICATION NUMBER: US/10/005,956
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/251,015
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/263,678
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/273,037
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1579
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 14
; LENGTH: 407
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-005-956-14

Query Match          100.0%; Score 114; DB 4; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFAEASMAAFNTVNF 23
    |||||
Db 68 NYFLVNLAFAEASMAAFNTVNF 90

RESULT 13
US-10-005-956-16
; Sequence 16, Application US/10005956
; Publication No. US20030113726A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: D0053NP
; CURRENT APPLICATION NUMBER: US/10/005,956
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/251,015
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/263,678
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/273,037
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1579
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 16
; LENGTH: 407
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-005-956-16

Query Match          100.0%; Score 114; DB 4; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYFLVNLAFAEASMAAFNTVNF 23
    |||||
Db 68 NYFLVNLAFAEASMAAFNTVNF 90

RESULT 14
US-10-005-956-18
; Sequence 18, Application US/10005956
; Publication No. US20030113726A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: D0053NP
; CURRENT APPLICATION NUMBER: US/10/005,956
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/251,015
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/263,678
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/273,037
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1579
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 18
; LENGTH: 407
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-005-956-18
```

Query Match 100.0%; Score 114; DB 4; Length 407;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPLVNLAFAPASMAAFTVNF 23
 |||||
 Db 68 NYPLVNLAFAPASMAAFTVNF 90

RESULT 15

US-10-005-956-20
 ; Sequence 20, Application US/10005956
 ; Publication No. US20030113726A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
 ; FILE REFERENCE: D0053NP
 ; CURRENT APPLICATION NUMBER: US/10/005,956
 ; PRIOR FILING DATE: 2001-12-03
 ; PRIOR APPLICATION NUMBER: 60/251,015
 ; PRIOR FILING DATE: 2000-12-04
 ; PRIOR APPLICATION NUMBER: 60/263,678
 ; PRIOR FILING DATE: 2001-01-23
 ; PRIOR APPLICATION NUMBER: 60/273,037
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 1579
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 20
 ; LENGTH: 407
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-005-956-20

Query Match 100.0%; Score 114; DB 4; Length 407;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYPLVNLAFAPASMAAFTVNF 23
 |||||
 Db 68 NYPLVNLAFAPASMAAFTVNF 90

Search completed: January 23, 2006, 12:09:00
 Job time : 52.8297 secs

This page blank (except for)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:53:39 ; Search time 3.21397 Seconds
(without alignments)
72.521 Million cell updates/sec

Title: US-10-501-838a-24

Perfect score: 114
Sequence: 1 NYFLVNLAFBASMALFTVNVNF 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/protdata/2/pubppaa/US08_NEW_PUB pep.*
2: /cgn2_6/protdata/2/pubppaa/US06_NEW_PUB pep.*
3: /cgn2_6/protdata/2/pubppaa/US07_NEW_PUB pep.*
4: /cgn2_6/protdata/2/pubppaa/PCT_NEW_PUB pep.*
5: /cgn2_6/protdata/2/pubppaa/US03_NEW_PUB pep.*
6: /cgn2_6/protdata/2/pubppaa/US10_NEW_PUB pep.*
7: /cgn2_6/protdata/2/pubppaa/US11_NEW_PUB pep.*
8: /cgn2_6/protdata/2/pubppaa/US60_NEW_PUB pep.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	114	100.0	407	US-11-127-877-56	Sequence 56, Appl
2	101	88.6	23	US-11-127-877-571	Sequence 571, Appl
3	47	41.2	269	US-11-151-482-5	Sequence 5, Appl
4	44	38.6	445	US-11-115-564-1	Sequence 1, Appl
5	44	38.6	445	US-11-115-564-2	Sequence 2, Appl
6	44	38.6	445	US-11-115-564-3	Sequence 3, Appl
7	42.5	37.3	590	US-11-124-368A-183	Sequence 183, Appl
8	42.5	37.3	590	US-11-127-877-54	Sequence 54, Appl
9	42	36.8	95	US-10-521-162-37	Sequence 37, Appl
10	42	36.8	480	US-10-521-162-40	Sequence 40, Appl
11	41.5	36.4	466	US-11-127-877-41	Sequence 41, Appl
12	41	36.0	231	US-10-689-742-132	Sequence 132, Appl
13	41	36.0	315	US-10-467-657-4960	Sequence 4960, Appl
14	41	36.0	481	US-11-090-439-16	Sequence 16, Appl
15	40	35.1	30	US-10-901-576-10	Sequence 10, Appl
16	40	35.1	31	US-10-901-576-9	Sequence 9, Appl
17	40	35.1	890	US-11-106-623-28	Sequence 28, Appl
18	39	34.2	272	US-10-793-626-3198	Sequence 3198, Appl
19	39	34.2	419	US-11-067-884-8	Sequence 8, Appl
20	39	34.2	439	US-11-127-877-51	Sequence 51, Appl
21	39	34.2	434	US-11-156-084-229	Sequence 229, Appl
22	39	34.2	466	US-11-127-877-50	Sequence 50, Appl
23	39	34.2	471	US-10-995-561-901	Sequence 901, Appl
24	38.5	33.8	310	US-11-212-443-192	Sequence 192, Appl
25	38.5	33.8	319	US-11-212-443-68	Sequence 68, Appl

25	38.5	33.8	330	7	US-11-212-443-193	Sequence 193, Appl
27	38.5	33.8	540	7	US-11-212-443-167	Sequence 167, Appl
28	38.5	33.8	540	7	US-11-212-443-168	Sequence 168, Appl
29	38.5	33.8	595	7	US-11-212-443-72	Sequence 72, Appl
30	38.5	33.8	619	7	US-11-212-443-70	Sequence 70, Appl
31	38.5	33.8	640	7	US-11-212-443-20	Sequence 20, Appl
32	38.5	33.8	761	7	US-11-212-443-22	Sequence 22, Appl
33	38.5	33.8	723	7	US-11-127-877-556	Sequence 556, Appl
34	38	33.3	394	7	US-11-074-176-310	Sequence 310, Appl
35	38	33.3	409	6	US-10-627-633-4	Sequence 4, Appl
36	38	33.3	414	7	US-11-074-176-46	Sequence 46, Appl
37	38	33.3	415	6	US-10-627-633-2	Sequence 2, Appl
38	38	33.3	424	6	US-10-467-657-2678	Sequence 2678, Appl
39	38	33.3	456	6	US-10-966-483-34	Sequence 34, Appl
40	38	33.3	456	7	US-11-021-441-18	Sequence 18, Appl
41	38	33.3	479	6	US-10-966-483-36	Sequence 36, Appl
42	38	33.3	479	7	US-11-021-441-20	Sequence 20, Appl
43	38	33.3	490	6	US-10-966-483-42	Sequence 42, Appl
44	38	33.3	490	7	US-11-021-441-26	Sequence 26, Appl
45	38	33.3	497	6	US-10-966-483-38	Sequence 38, Appl

ALIGNMENTS

```
RESULT 1
US-11-127-877-56
; Sequence 56, Application US/11127877
; Publication No. /US20050287565A1
; GENERAL INFORMATION:
; APPLICANT: Merchiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koentraad F. F.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27, 800-B USA
; CURRENT APPLICATION NUMBER: US/11/127,877
; PRIOR FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 590
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 56
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-877-56

Query Match          100.0%; Score 114; DB 7; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYFLVNLAFBASMALFTVNVNF 23
Db 68 NYFLVNLAFBASMALFTVNVNF 90

RESULT 2
US-11-127-877-571
; Sequence 571, Application US/11127877
; Publication No. US20050287565A1
; GENERAL INFORMATION:
; APPLICANT: Merchiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spittaels, Koentraad F. F.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27, 800-B USA
; CURRENT APPLICATION NUMBER: US/11/127,877
```

```

; CURRENT FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 590
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 571
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-877-571

Query Match      88.6%; Score 101; DB 7; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.7e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      3 FLVNLAFASMAAFNTVNF 23
Db      1 FLVNLAFASMAAFNTVNF 21

RESULT 3
US-11-151-482-5
; Sequence 5, Application US/11151482
; Publication No. US20060002919A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 15625 Receptor, A Novel G-Protein Coupled Receptor
; FILE REFERENCE: 5800-13, 035800-171548
; CURRENT APPLICATION NUMBER: US/11/151,482
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US/09/187,134
; PRIOR FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 5
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Rhodopsin family
US-11-151-482-5

Query Match      41.2%; Score 47; DB 7; Length 269;
Best Local Similarity 52.9%; Pred. No. 0.96;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy      1 NYFLVNLAFASMAAF 17
Db      20 NYFLVNLAVDLFLSLP 36

RESULT 4
US-11-115-564-1
; Sequence 1, Application US/11115564
; Publication No. US20050267116A1
; GENERAL INFORMATION:
; APPLICANT: Peschke, Bernd
; APPLICANT: Hohlweg, Rolf
; TITLE OF INVENTION: SUBSTITUTED HEXAHYDROPYRROLO[1,2-a]PYRAZINES,
; TITLE OF INVENTION: OCTAHYDROPYRIDO[1,2-a]PYRAZINES AND
; FILE REFERENCE: 6483.200-US
; CURRENT APPLICATION NUMBER: US/11/115,564
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: US 60/387,047
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: Danish Application no. PA 2002 00863
; PRIOR FILING DATE: 2002-06-06
; NUMBER OF SEQ ID NOS: 3
```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 1
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-115-564-1

Query Match      38.6%; Score 44; DB 7; Length 445;
Best Local Similarity 47.1%; Pred. No. 5.4;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Cy      1 NYFLVNLAFASMAAF 17
Db      70 NYFLNLALSDPLVGAF 86

RESULT 5
US-11-115-564-2
; Sequence 2, Application US/11115564
; Publication No. US20050267116A1
; GENERAL INFORMATION:
; APPLICANT: Peschke, Bernd
; APPLICANT: Hohlweg, Rolf
; TITLE OF INVENTION: SUBSTITUTED HEXAHYDROPYRROLO[1,2-a]PYRAZINES,
; TITLE OF INVENTION: OCTAHYDROPYRIDO[1,2-a]PYRAZINES AND
; FILE REFERENCE: 6483.200-US
; CURRENT APPLICATION NUMBER: US/11/115,564
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: US 60/387,047
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: Danish Application no. PA 2002 00863
; PRIOR FILING DATE: 2002-06-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Monkey
US-11-115-564-2

Query Match      38.6%; Score 44; DB 7; Length 445;
Best Local Similarity 47.1%; Pred. No. 5.4;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Cy      1 NYFLVNLAFASMAAF 17
Db      70 NYFLNLALSDPLVGAF 86

RESULT 6
US-11-115-564-3
; Sequence 3, Application US/11115564
; Publication No. US20050267116A1
; GENERAL INFORMATION:
; APPLICANT: Peschke, Bernd
; APPLICANT: Hohlweg, Rolf
; TITLE OF INVENTION: SUBSTITUTED HEXAHYDROPYRROLO[1,2-a]PYRAZINES,
; TITLE OF INVENTION: OCTAHYDROPYRIDO[1,2-a]PYRAZINES AND
; FILE REFERENCE: 6483.200-US
; CURRENT APPLICATION NUMBER: US/11/115,564
; CURRENT FILING DATE: 2005-04-27
; PRIOR APPLICATION NUMBER: US 60/387,047
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: Danish Application no. PA 2002 00863
; PRIOR FILING DATE: 2002-06-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 3
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Rat
```

RESULT 11
US-11-127-877-41

; Sequence 41, Application US/11127877
; Publication No. US20050287565A1
; GENERAL INFORMATION:
; APPLICANT: Merchiers, Pascal G.
; APPLICANT: Hoffmann, Marcel
; APPLICANT: Spitznagel, Koentraad F. F.
; APPLICANT: Laenen, Wendy
; TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
; TITLE OF INVENTION: Amyloid-Beta Protein Production
; FILE REFERENCE: P27, 800-B USA
; CURRENT APPLICATION NUMBER: US/11/127,877
; PRIOR FILING DATE: 2005-05-12
; PRIOR APPLICATION NUMBER: 60/570,352
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: 60/603,948
; PRIOR FILING DATE: 2004-08-24
; NUMBER OF SEQ ID NOS: 590
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 41
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-127-877-41

Query Match 36.4%; Score 41.5; DB 7; Length 466;
Best Local Similarity 33.3%; Pred. No. 15;
Matches 10; Conservative 6; Mismatches 7; Indels 7; Gaps 1;

OY 1 NYPLVNLAFAPASMAAFNTVNF 23
||| : ||| :
Db 59 NYPLFSLACADLLIGVFAMNLYTYTVIGY 88

RESULT 12
US-10-689-742-132
; Sequence 132, Application US/10689742
; Publication No. US20050250180A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M
; APPLICANT: Lavallie, Edward R
; APPLICANT: Racle, Lisa A
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 00766, 000091.10
; CURRENT APPLICATION NUMBER: US/10/689,742
; CURRENT FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: 09/746,783
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 132
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-742-132

Query Match 36.0%; Score 41; DB 6; Length 231;
Best Local Similarity 34.8%; Pred. No. 7.8;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

OY 1 NYPLVNLAFAPASMAAFNTVNF 23
||| : ||| :
Db 146 NYPLVAVTSLSSQIQYMTVTVNF 168

RESULT 13
US-10-467-657-4960
; Sequence 4960, Application US/10467657
; Publication No. US20050260581A1

; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4960
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4960

Query Match 36.0%; Score 41; DB 6; Length 315;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 VNLAFAESMAAFNTV 20
||| : ||| :
Db 266 VNLAFAEGYAAAFNNL 281

RESULT 14
US-11-090-439-16
; Sequence 16, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberculous Sclerosis Null
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-16

Query Match 36.0%; Score 41; DB 7; Length 481;
Best Local Similarity 47.1%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 NYPLVNLAFAPASMAAF 17
||| : ||| :
Db 90 NYPLMSLAVADLVGLF 106

RESULT 15
US-10-901-576-10
; Sequence 10, Application US/10901576
; Publication No. US2006009913A1
; GENERAL INFORMATION:
; APPLICANT: Trabantino, Rene J.
; APPLICANT: Valdehl, Nagarajan
; APPLICANT: Hall, Spencer B.
; APPLICANT: Goddard, William A.
; APPLICANT: Floriano, Wely
; TITLE OF INVENTION: SYSTEM AND METHODS FOR PREDICTING
; TITLE OF INVENTION: TRANSMEMBRANE DOMAINS IN MEMBRANE PROTEINS AND MINING THE
; TITLE OF INVENTION: GENOME FOR RECOGNIZING G-PROTEIN COUPLED RECEPTORS

FILE REFERENCE: CTCH-P01-023
 CURRENT APPLICATION NUMBER: US/10/901,576
 CURRENT FILING DATE: 2004-07-29
 PRIOR APPLICATION NUMBER: US 60/491,334
 PRIOR FILING DATE: 2003-07-29
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 10
 LENGTH: 30
 TYPE: prt
 ORGANISM: Bos taurus
 US-10-901-576-10

Query Match 35.1%; Score 40; DB 6; Length 30;
 Best Local Similarity 43.5%; Pred. No. 1.1;
 Matches 10; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 NYFLVNLPAPASM--AAPNTTV 21
 |||:|||||:
 DB 3 NYLLNLAVADLFMVFGFTTL 25

Search completed: January 23, 2006, 12:09:42
 Job time : 3.21397 secs

The Negro Bank (ustrio)
(Olden) yubig oden: 944

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:32:56 ; Search time 13.1004 seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-25

Perfect score: 119

Sequence: 1 MNYYDITLALGICQSARLVQOLA 24

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	114	95.8	208	2	AH0199	conserved hypothet
2	114	95.8	213	2	S19211	ycfc protein - Esc
3	114	95.8	213	2	D90829	hypothetical prote
4	114	95.8	213	2	B85687	hypothetical prote
5	111	93.3	215	2	AB0647	conserved hypothet
6	95	79.8	205	2	I64155	hypothetical prote
7	80	67.2	211	2	D84860	hypothetical prote
8	72	60.5	205	2	B82337	conserved hypothet
9	72	60.5	205	2	I57957	neurokinin 2 recep
10	48	40.3	384	2	S20303	neurokinin 2 recep
11	48	40.3	390	2	A36737	neurokinin 2 recep
12	47	39.5	144	2	B91207	heat shock protein
13	47	39.5	144	2	G65170	heat shock protein
14	47	39.5	144	2	D86053	neurokinin 2 recep
15	47	39.5	384	1	S00516	neurokinin 2 recep
16	47	39.5	398	1	JQ1059	neurokinin 2 recep
17	47	39.5	402	2	I56595	neurokinin 2 recep
18	46	38.7	142	2	AF0961	heat shock protein
19	46	38.7	154	2	AC0496	heat shock protein
20	46	38.7	317	2	A44156	lipopolysaccharide
21	46	38.7	335	2	S70671	lipopolysaccharide
22	45	37.8	188	2	JC1065	coat protein - bee
23	45	37.8	947	2	AH0891	adenyl-transferase
24	44	37.0	204	2	E82680	conserved hypothet
25	44	37.0	280	2	AI0326	cel operon repres
26	44	37.0	375	1	NREBCD	ribonuclease III (
27	44	37.0	375	2	A99943	RNase D, processes
28	44	37.0	375	2	E85791	RNase D, processes
29	44	37.0	433	2	T25946	hypothetical prote

30	43	36.1	282	2	E71543	hypothetical prote
31	43	36.1	299	2	B81698	conserved hypothet
32	43	36.1	411	2	T00567	hypothetical prote
33	43	36.1	514	2	B96653	hypothetical prote
34	43	36.1	759	2	D70422	cellulose synthase
35	43	36.1	847	2	G95843	conserved hypothet
36	42	35.3	98	2	T22503	hypothetical prote
37	42	35.3	145	2	S41221	hypothetical prote
38	42	35.3	224	2	G82780	conserved hypothet
39	42	35.3	273	2	AC3277	transposase BME102
40	42	35.3	336	2	G71954	lipopolysaccharide
41	42	35.3	340	2	G64554	lipopolysaccharide
42	42	35.3	648	1	NPVZCP	nucleoside-triphos
43	41.5	34.9	510	2	D6741	hypothetical prote
44	41.5	34.9	1215	2	T43916	chitinase A (import
45	41.5	34.9	1850	2	AC1917	serine/threonine K

ALIGNMENTS

RESULT 1
AH0199
Conserved hypothetical protein YP01637 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence revision 02-Nov-2001 #text_change 09-Jul-2004
R/Parthill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11585360
A/Accession: AH0199
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-208 <KUR>
A/Cross-references: UNIPROT:Q8ZF06; UNIPARC:UPI00000CD828; GB:AL590842; PIDN:CAC90459.1
A/Genetics:
A/Gene: YP01637

Query Match 95.8%; Score 114; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 6e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 MNYYDITLALGICQSARLVQOLA 24
DB 4 NYDITLALGICQSARLVQOLA 26

RESULT 2

ycfc protein - Escherichia coli (strain K-12)

C/Species: Escherichia coli

C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text_change 09-Jul-2004

C/Accession: S19211; A64858

R/Green, S.M.; Drablie, W.T.
submitted to the EMBL Data Library, May 1991

A/Description: Molecular analysis of the purB-phop region of Escherichia coli K12.

A/Reference number: S19210

A/Accession: S19211

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-213 <GRB>

A/Cross-references: UNIPROT:P25746; UNIPARC:UPI000013A673; EMBL:X59307; NID:G42582; PID:

R/Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C

.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: A64858
A/Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-213 <BLAT>
A:Cross-references: UNIPARC:UPI000013A673; GB:AE000213; GB:U00096; NID:g1787371; PIDN:AA
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yfcC
A:start codon: GTG

Query Match 95.8%; Score 114; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 6.1e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDYITLALAGICQSARLVQQLA 24
Db 4 NYDYITLALAGICQSARLVQQLA 26

RESULT 3
D90829
hypotheetical protein ECA1604 [imported] - Escherichia coli (strain O157:H7, substrain R1
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D90829
R:Havashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sawawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A96929; MUID:21156231; PMID:11258796
A:Accession: D90829
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <HAY>
A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:BA000007; PIDN:BA035027.1;
A:Experimental source: strain O157:H7, substrain RMD 050952
C:Genetics:
A:Gene: ECA1604

Query Match 95.8%; Score 114; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 6.1e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDYITLALAGICQSARLVQQLA 24
Db 4 NYDYITLALAGICQSARLVQQLA 26

RESULT 4
B85687
hypotheetical protein yfcC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85687
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, B.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:AE005174; NID:g12514786; F
C:Genetics:
A:Experimental source: strain O157:H7, substrain EDL933
A:Gene: yfcC

Query Match 95.8%; Score 114; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 6.1e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDYITLALAGICQSARLVQQLA 24
Db 4 NYDYITLALAGICQSARLVQQLA 26

RESULT 5
AB0647
conserved hypotheetical protein STY1273 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhimurium
A>Note: this species has also been called Salmonella typhimurium
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0647
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0647
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-215 <PAR>
A:Cross-references: UNIPARC:UPI0000059F77; GB:AL513382; PIDN:CAD08357.1; PID:g16502402;
C:Genetics:
A:Gene: STY1273

Query Match 93.3%; Score 111; DB 2; Length 215;
Best Local Similarity 95.7%; Pred. No. 1.9e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDYITLALAGICQSARLVQQLA 24
Db 6 NYDYITLALAGICQSARLVQQLA 28

RESULT 6
164155
hypotheetical protein H10638 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: 164155
R:Flaischman, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
; D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fritchman, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A6400; MUID:95350630; PMID:7542800
A:Accession: 164155
A:Molecule type: DNA
A:Residues: 1-205 <TTGR>
A:Cross-references: UNIPROT:P44796; UNIPARC:UPI000013A674; GB:U32747; GB:LA2023; NID:g15
A>Note: best homolog was a hypotheetical protein from Escherichia coli

Query Match 79.8%; Score 95; DB 2; Length 205;
Best Local Similarity 78.3%; Pred. No. 7.4e-08;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 NYDYITLALAGICQSARLVQQLA 24
Db 3 NYHDIIVLALAGVCSAKLVHQLA 25

RESULT 7
DB4960
hypotheetical protein [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: DB4960
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A>Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: DB4960

A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:GeneID: EC64626
A:Gene: EC64626
C:Superfamily: alpha-crystallin-related small heat shock protein

Query Match 39.5%; Score 47; DB 2; Length 144;
Best Local Similarity 55.0%; Pred. No. 3.6;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQARLVQ 21
DB 45 NHYRITLALAGFRQEDLEIQ 64

RESULT 13
G65170
heat shock protein hspb - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: G65170; B45245
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: G65170
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-144 <BLAT>
A:Cross-references: UNIPROT:O8XC04; UNIPARC:UPI00000D0522C; GB:AE000445; GB:U00096; NID:G
R:Allen, S.P.; Polazzi, J.O.; Gierse, J.K.; Easton, A.M.
J. Bacteriol. 174, 6938-6947, 1992
A:Title: Two novel heat shock genes encoding proteins produced in response to heterolog
A:Reference number: A45245; MUID:93015757; PMID:1356969
A:Accession: B45245
A:Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 3-144 <ALU>
A:Cross-references: UNIPARC:UPI000012D125; GB:M94104; NID:9147366; PIDN:AAA24425.1; PID:
A:Note: sequence extracted from NCBI backbone (NCBIN:116883; NCBIPI:116885)
C:GeneID:
A:Gene: hsp15
C:Superfamily: alpha-crystallin-related small heat shock protein

Query Match 39.5%; Score 47; DB 2; Length 144;
Best Local Similarity 55.0%; Pred. No. 3.6;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQARLVQ 21
DB 45 NHYRITLALAGFRQEDLEIQ 64

RESULT 14
D86053
heat shock protein (imported) - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D86053
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lm, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A65480; MUID:21074935; PMID:11206551
A:Accession: D86053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <STO>
A:Cross-references: UNIPROT:O8XC04; UNIPARC:UPI00000D0522C; GB:AE005174; NID:912518527; F
C:GeneID:
A:Gene: hspb

C:Superfamily: alpha-crystallin-related small heat shock protein

Query Match 39.5%; Score 47; DB 2; Length 144;
Best Local Similarity 55.0%; Pred. No. 3.6;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQARLVQ 21
DB 45 NHYRITLALAGFRQEDLEIQ 64

RESULT 15
S00516
neurokinin 2 receptor - bovine
N:Alternate names: neurokinin A receptor; NK-2 receptor; substance K receptor (SKR)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S00516
R:Mausu, Y.; Nakayama, K.; Tamaki, H.; Harada, Y.; Kuno, M.; Nakanishi, S.
Nature 329, 836-838, 1997
A:Title: cDNA cloning of bovine substance-K receptor through oocyte expression system.
A:Reference number: S00516; MUID:88039072; PMID:2823146
A:Accession: S00516
A:Molecule type: mRNA
A:Residues: 1-384 <MAS>
A:Cross-references: UNIPROT:P05363; UNIPARC:UPI00001301EE; EMBL:X06295; NID:9746; PIDN:C
C:Superfamily: neurokinin 1 receptor
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane pr
F:33-59/Domain: transmembrane #status predicted <TM1>
F:70-93/Domain: transmembrane #status predicted <TM2>
F:111-159/Domain: transmembrane #status predicted <TM3>
F:150-169/Domain: transmembrane #status predicted <TM4>
F:193-222/Domain: transmembrane #status predicted <TM5>
F:252-275/Domain: transmembrane #status predicted <TM6>
F:286-310/Domain: transmembrane #status predicted <TM7>
F:11/Binding site: carbohydrate (Aan) (covalent) #status predicted
F:106-181/Dissulfide bonds: #status predicted

Query Match 39.5%; Score 47; DB 1; Length 384;
Best Local Similarity 50.0%; Pred. No. 9.3;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQSA 17
DB 69 NYFVNLALADLCMAA 84

Search completed: January 23, 2006, 10:43:09
Job time : 13.1004 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:31:19 ; Search time 76.5066 Seconds
(without alignments)
221.323 Million cell updates/sec

Title: US-10-501-838a-25
Perfect score: 119
Sequence: 1 MNYDYITLALAGICQSAARLVQOLA 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	95.8	208	1 Y1637	YERPE
2	114	95.8	208	2 Q669Q3	YERPS
3	114	95.8	213	1 YCFC	ECOS7
4	114	95.8	213	1 YCFC	ECOL6
5	114	95.8	213	1 YCFC	ECOLI
6	114	95.8	213	1 YCFC	SHITL
7	111	93.3	213	1 YCFC	SALTI
8	111	93.3	213	1 YCFC	SALTI
9	111	93.3	213	2 Q5PMJ3	SALPA
10	107	89.9	215	2 Q5PMJ3	SALPA
11	106	89.1	204	2 Q65VY5	MANSM
12	105	88.2	212	2 Q6D4B8	ERWCT
13	103	86.6	212	2 Y1850	PASWU
14	97	79.8	208	1 Y2805	PHOHL
15	95	79.8	205	1 Y638	HARIN
16	95	79.8	205	2 Q4QW59	HAE18
17	82	68.9	215	1 Y1650	HAEDU
18	80	67.2	211	1 Y262	BUCAI
19	74	62.2	217	1 Y243	BUCBP
20	72	60.5	205	1 Y1127	VIBCB
21	72	60.5	205	1 Y1129	VIBBP
22	70	58.8	205	1 Y1347	VIBBV
23	70	58.8	205	1 Y2927	VIBVU
24	69	58.0	205	2 Q5E3W6	VIBF1
25	62	52.1	205	2 Q6L1T9	PHOPR
26	56	47.1	468	2 Q9VGQ1	DROME
27	53	44.5	372	2 Q7PSM6	ANOGA
28	50	42.0	150	1 IBPB	BRWCT
29	49	40.3	391	2 Q5GVQ3	XANOR
30	48	40.3	197	2 Q98ER3	RHILQ
31	48	40.3	207	2 Q60CA8	METCA

32	48	40.3	216	2 Q6C8B5	YARLI
33	48	40.3	266	2 Q7S4H4	NEURC
34	48	40.3	274	2 Q5CSB5	CRYPR
35	48	40.3	274	2 Q5CPS8	CRYHO
36	48	40.3	384	1 NK2R	MESAV
37	48	40.3	384	1 NK2R	MOUSE
38	48	40.3	384	2 Q5DUB0	MERUN
39	48	40.3	384	2 Q8BZV9	MOUSE
40	48	40.3	390	1 NK2R	RAT
41	48	40.3	473	2 Q5V518	HALMA
42	48	40.3	599	2 Q726U4	DESVH
43	48	40.3	3223	2 Q5KSU8	ECOLI
44	48	40.3	3223	2 Q8L1L9	ECOLI
45	48	40.3	3223	2 Q8VNR6	ECOLI

ALIGNMENTS

RESULT 1
Y1637_YERPE STANDARD; PRT; 208 AA.
ID Y1637_YERPE
AC Q8ZFQ6; Q8DQ01;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein YP01637/Y1798/Y1767.
GN Ordered locus names=YP01637, Y1798, Y1767;
OS *Yersinia pestis*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Yersinia*.
OX NCBI_TaxID=632;
ON (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titchall R.W., Houlden M.T.G.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Farraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Ovelton P.C.F., Quail M.A., Rutherford K.M.,
RA Stimpson M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,
RA "Genome sequence of *Yersinia pestis*, the causative agent of plague",
Nature 413:523-527 (2001).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burdett V., Plunkett G., III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.B., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
RA Perry R.D.,
RT "Genome sequence of *Yersinia pestis* KIM",
J. Bacteriol. 184:4601-4611 (2002).
[3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cai B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.,
RT "Complete genome sequence of *Yersinia pestis* strain 91001, an isolate
avirulent to humans",
DNA Res. 11:1179-1197 (2004).
CC - SIMILARITY: Belongs to the YP0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its

```

CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; AJ414149; CAC90459.1; -; Genomic_DNA.
DR EMBL; AE013782; AAM85366.1; ALT_INIT; Genomic_DNA.
DR EMBL; AE017133; AAS61994.1; -; Genomic_DNA.
DR PIR; AH0199; AH0199.
DR SMR; Q82F06; 2-207.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
FT CONFLICT 1 M -> MIV (in Ref. 2).
SQ SEQUENCE 208 AA; 22730 MW; DQCC2B3BDE499723 CRC64;

Query Match 95.8%; Score 114; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.4e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 NYDITLALAGICQSARLVQQLA 24
Db 4 NYDITLALAGICQSARLVQQLA 26

RESULT 2
069903_YERPS PRELIMINARY; PRT; 208 AA.
AC Q66903;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=YPRB2431;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype 1;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chai P.S.G., Camfield E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Himebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenaï-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR SML; BX936398; CAH21669.1; -; Genomic_DNA.
DR SMR; Q66903; 2-207.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 208 AA; 22730 MW; DQCC2B3BDE499723 CRC64;

Query Match 95.8%; Score 114; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.4e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 NYDITLALAGICQSARLVQQLA 24
Db 4 NYDITLALAGICQSARLVQQLA 26

RESULT 3
YCFC_ECOL57 STANDARD; PRT; 213 AA.
ID YCFC_ECOL57
AC Q8Y736;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.


```

```

GN Name=ycfC; OrderedLocustNames=z1861, ECol604;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta R.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC -1 SIMILARITY: Belongs to the UPF0274 family.
CC CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; AE005174; AAC55958.1; -; Genomic_DNA.
DR EMBL; BA000007; BAB35027.1; -; Genomic_DNA.
DR PIR; B85687; B85687.
DR PIR; D90829; D90829.
DR SMR; Q8Y736; 2-213.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22947 MW; E26EP9678C3844E2 CRC64;

Query Match 95.8%; Score 114; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 7.6e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 NYDITLALAGICQSARLVQQLA 24
Db 4 NYDITLALAGICQSARLVQQLA 26

RESULT 4
YCFC_ECOL6 STANDARD; PRT; 213 AA.
ID YCFC_ECOL6
AC Q8F1B7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocustNames=c1511;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;


```


RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Raeko D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic *Escherichia coli*,"
 RT Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL, AB016759; AAN79980.1; ALT_INIT; Genomic_DNA.
 DR SMR, Q8P1B7; 2-213.
 DR HAMAP, MF_00695; -, 1.
 DR InterPro, IPR007451; DUF489.
 DR Pfam, PF04356; DUF489; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 213 AA; 22934 MW; 71C8F96F8BA37BB CRC64;
 Query Match 95.8%; Score 114; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.6e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 NYDITLALAGICQSARLVQOLA 24
 Db 4 NYDITLALAGICQSARLVQOLA 26
 RESULT 5
 YCFC_ECOLI STANDARD; PRT; 213 AA.
 AC P25746;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC (ORF-23).
 GN Name:ycfC; Ordered locus names=b1132;
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OX NCBI_TaxID=562;
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K12;
 RA Green S.M., Drabble W.T.;
 RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K12;
 RX MEDLINE=92104952; PubMed=1729205;
 RT He B., Smith J.M., Zaikin H.;
 RT "Escherichia coli purr gene: cloning, nucleotide sequence, and
 RT regulation by purr,"
 RT J. Bacteriol. 174:130-136(1992).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kittpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12,"
 RT Science 277:1453-1474(1997).
 RL [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Ohnita T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horuchi T.;
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map,"
 RT DNA Res. 3:137-155(1996).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to
 CC frameshifts.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL, X59307; CAA41995.1; -, Genomic_DNA.
 DR EMBL, M74924; -, NOT ANNOTATED CDS; Genomic_DNA.
 DR EMBL, U00096; AAC74216.1; -, Genomic_DNA.
 DR EMBL, D90748; BAA35954.1; -, Genomic_DNA.
 DR EMBL, D90749; BAA35963.1; -, Genomic_DNA.
 DR PIR, S19211; S19211.
 DR PDB, 1QZ4; X-ray; A=2-213.
 DR PDB, 1SD1; X-ray; A=2-213.
 DR ECHOBASE; EBI321; -.
 DR EcoGene; EG1345; ycfC.
 DR HAMAP, MF_00695; -, 1.
 DR InterPro, IPR007451; DUF489.
 DR Pfam, PF04356; DUF489; 1.
 KW 3D structure; Complete proteome; Hypothetical protein.
 SQ SEQUENCE 213 AA; 22948 MW; E26F9698C3CE842 CRC64;
 Query Match 95.8%; Score 114; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.6e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 NYDITLALAGICQSARLVQOLA 24
 Db 4 NYDITLALAGICQSARLVQOLA 26
 RESULT 6
 YCFC_SHIFL STANDARD; PRT; 213 AA.
 AC O83LF8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC.
 GN Name:ycfC; Ordered locus names=SP1151, S1234;
 OS *Shigella flexneri*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Shigella*.
 OX NCBI_TaxID=623;
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu Y., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
 RT through comparison with genomes of *Escherichia coli* K12 and O157,"
 RT Nucleic Acids Res. 30:4432-4441(2002).
 RL [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 DOI=10.1128/JAI.71.5.2775-2786.2003;

```

CC      RA Mel J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA      Rounhell G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA      Mau B., Perma N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA      Schwartz D.C., Blattner F.R.;
RT      "Complete genome sequence and comparative genomics of Shigella
RT      flexneri serotype 2a strain 2457T."
RL      Infect. Immun. 71:2775-2786(2003).
CC      CC      -1- SIMILARITY: Belongs to the UPF0274 family.
CC      -----
CC      CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      CC      the European Bioinformatics Institute. There are no restrictions on its
CC      CC      use as long as its content is in no way modified and this statement is not
CC      CC      removed.
CC      CC      -----
CC      DR EMBL; AE005674; AAN42768.1; ALT INIT; Genomic_DNA.
CC      DR EMBL; AE016982; AAP16657.1; -; Genomic_DNA.
CC      DR SRR; Q83LF8; 2-213.
CC      DR HAMAP; MF_00695; -; 1.
CC      DR InterPro; IPR007451; DUF489.
CC      DR Pfam; PF04356; DUF489; 1.
CC      KW Complete proteome; Hypothetical protein.
CC      SEQUENCE 213 AA; 22894 MW; 08DBECC98C29FB4D CRC64;
CC
CC      Query Match 95.8%; Score 114; DB 1; Length 213;
CC      Best Local Similarity 100.0%; Pred. No. 7.6e-10;
CC      Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC      QY 2 NYDITLALAGICQSARLVQOLA 24
CC      Db 4 NYDITLALAGICQSARLVQOLA 26
CC
CC      RESULT 7
CC      YCFC_SALTI
CC      ID YCFC_SALTI STANDARD; PRT; 213 AA.
CC      AC Q827H0;
CC      DT 10-OCT-2003 (Rel. 42, Created)
CC      DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC      DT 10-MAY-2005 (Rel. 47, Last annotation update)
CC      DE Hypothetical UPF0274 protein YCFC.
CC      GN Name=ycfc; Ordered locus names=STRY1273, t1687;
CC      OS Salmonella typhi.
CC      OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC      OC Enterobacteriaceae; Salmonella.
CC      OX NCBI_TaxID=601;
CC      [1]
CC      RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC      RC STRAIN=CT18;
CC      RC MEDLINE=21534947; PubMed=11677808; DOI=10.1038/35101607;
CC      RA Parshall J., Dougan G., James K.D., Thomson S.R., Pickard D., Wain J.,
CC      RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
CC      RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
CC      RA Croft A., Davies P., Davies R.M., Dowd L., White N., Farrar J.,
CC      RA Feltham T., Hamlin T., Haque A., Hien T.T., Holtz S., Jorgensen K.,
CC      RA Krogh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
CC      RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
CC      RA Whitehead S., Barrett B.G.;
CC      RT "Complete genome sequence of a multiple drug resistant Salmonella
CC      RT enterica serovar Typhi CT18."
CC      RL Nature 413:848-852(2001).
CC      [2]
CC      RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC      RC STRAIN=Ty2 / ATCC 700931;
CC      RC MEDLINE=22531367; PubMed=12644504;
CC      RC DOI=10.1128/JB.185.7.2330-2337.2003;
CC      RA Deng W., Lyou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
CC      RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
CC      RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
CC      RT J. Bacteriol. 185:2330-2337(2003).
CC      -1- SIMILARITY: Belongs to the UPF0274 family.
CC

```

```

CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC
CC      EMBL; AL627269; CAD08357.1; ALT_INT; Genomic DNA.
DR      EMBL; AB016839; AA069312.1; ALT_INT; Genomic_DNA.
DR      SMK; Q827H0; 2-213.
DR      HAMAP; MF_00695; -; 1.
DR      InterPro; IPR007451; DUF489.
DR      Pfam; PF04356; DUF489; 1.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE 213 AA; 22942 MW; 63192D1C8066ED58 CRC64;

Query Match          93.3%; Score 111; DB 1; Length 213;
Best Local Similarity 95.7%; Pred. No. 2.3e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 NYDYITLALGICQSARLVQOLA 24
      |||||:|||||:|||||:|||||:
DB      4 NYDYITLALSGICQSARLVQOLA 26

RESULT 8
YCFC_SALTY          STANDARD;      PRT;      213 AA.
ID      YCFC_SALTY
AC      Q8ZP25;
DT      10-OCT-2003 (Rel. 42, Created)
DT      10-OCT-2003 (Rel. 42, Last sequence update)
DT      10-MAY-2005 (Rel. 47, Last annotation update)
DE      Hypothetical UPF0274 protein yCfc.
DN      Name=yCfc; OrderedLocusNames=STM1233;
GN      Salmonella typhimurium.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxID=602;
RX      [1]
      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=L72 / SGSC1412 / ATCC 700720;
RX      MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA      McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA      Courtney L., Portwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA      Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA      Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA      Waterston R., Wilson R.K.;
RT      "Complete genome sequence of Salmonella enterica serovar Typhimurium
      L72."
RL      Nature 413:852-856(2001).
CC      -1- SIMILARITY: Belongs to the UPF0274 family.
CC
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC
CC      EMBL; AB008754; AAL20162.1; ALT_INT; Genomic_DNA.
DR      EMBL; AB008754; AAL20162.1; ALT_INT; Genomic_DNA.
DR      SMK; Q8ZP25; 2-213.
DR      STyGene; SG172727; yCFC.
DR      HAMAP; MF_00695; -; 1.
DR      InterPro; IPR007451; DUF489.
DR      Pfam; PF04356; DUF489; 1.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match          93.3%; Score 111; DB 1; Length 213;
Best Local Similarity 95.7%; Pred. No. 2.3e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 NYDYITLALGICQSARLVQOLA 24
      |||||:|||||:|||||:|||||:
DB      4 NYDYITLALSGICQSARLVQOLA 26

```

RESULT 9

OSPMJ3_SALPA

ID 05PMJ3_SALPA PRELIMINARY; PRT; 213 AA.

AC 05PMJ3_2005 (TREMBlrel. 29, Created)

DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)

DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)

DE Hypothetical protein yfcC.

GN Name=yfcC; OrderedlocusNames=SPA1617;

OS Salmonella paratyphi-a.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OC NCB1_Taxid=54388;

[1]

NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 9150;

RX PubMed=1531882; DOI=10.1038/ng1470;

RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,

RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozereky P., McEllan M.,

RA Hartline C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,

RA Leonard S., Strong C., Du F., Carter J., Kremicki C., Layman D.,

RA Delaunay K., Fronick C., Magrini V., Nhan M., Warren W., Flores L.,

RA Spleth J., Wilson R.K.;

RT "Comparison of genome degradation in Paratyphi A and Typhi, human-

restricted serovars of Salmonella enterica that cause typhoid.";

RL Nat. Genet. 36:1268-1274(2004).

DR InterPro: IPR007451; DUF489.

DR Pfam: PF04356; DUF489; 1.

SQ Complete proteome; Hypothetical protein.

Query Match 213 AA; 22916 MW; 63192DC807C5D58 CRC64;

Best Local Similarity 93.3%; Score 111; DB 2; Length 213;

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDDITLALAGCQSAKLVQQA 24

DB 4 NYDDITLALAGCQSAKLVQQA 26

RESULT 10

Q570C1_SALCH

ID Q570C1_SALCH PRELIMINARY; PRT; 215 AA.

AC Q570C1_2005 (TREMBlrel. 30, Created)

DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)

DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)

DE Hypothetical protein yfcC.

GN Name=yfcC; OrderedlocusNames=SC1184;

OS Salmonella cholerae-suis (Salmonella enterica).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OC NCB1_Taxid=591;

[1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=SC-867;

RX PubMed=15781495;

RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,

RA Wang H.-S., Lee Y.-S.;

RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a

highly invasive and resistant zoonotic pathogen.";

RT Nucleic Acids Res. 33:1690-1698(2005).

DR EMBL: AB017220; AA65090.1; -; Genomic DNA.

KM Complete proteome; Hypothetical protein.

SQ SEQUENCE 215 AA; 23157 MW; C153B20629FAE8BD CRC64;

Query Match 89.9%; Score 107; DB 2; Length 215;

Best Local Similarity 95.5%; Pred. No. 1e-08;

Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDDITLALAGCQSAKLVQQA 23

DB 6 NYDDITLALAGCQSAKLVQQA 27

RESULT 11

O65VY5_MANSM

ID O65VY5_MANSM PRELIMINARY; PRT; 204 AA.

AC O65VY5_2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)

DE Hypothetical protein.

GN OrderedlocusNames=MS0298;

OS Mannheimia succiniciproducens (strain MBE155E).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Mannheimia.

OC NCB1_Taxid=221986;

[1]

NUCLEOTIDE SEQUENCE.

RX PubMed=15378067; DOI=10.1038/nbt1010;

RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,

RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;

RT "The genome sequence of the capnophilic rumen bacterium Mannheimia

succiniciproducens.";

RL Nat. Biotechnol. 22:1275-1281(2004).

DR EMBL: AB016827; AAU36905.1; -; Genomic DNA.

DR InterPro: IPR007451; DUF489.

DR Pfam: PF04356; DUF489; 1.

KM Complete proteome; Hypothetical protein.

SQ SEQUENCE 204 AA; 23036 MW; 3D54A70A30C3CCB CRC64;

Query Match 89.1%; Score 106; DB 2; Length 204;

Best Local Similarity 87.0%; Pred. No. 1.4e-08;

Matches 20; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 NYDDITLALAGCQSAKLVQQA 24

DB 4 NYDDITLALAGCQSAKLVQQA 26

RESULT 12

Q6D4B8_ERWCT

ID Q6D4B8_ERWCT PRELIMINARY; PRT; 212 AA.

AC Q6D4B8_2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)

DE Hypothetical protein.

GN OrderedlocusNames=ECA2443;

OS Erwina carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Pectobacterium.

OC NCB1_Taxid=29471;

[1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=SCRI 1043 / ATCC BAA-672;

RX PubMed=15263089; DOI=10.1073/pnas.0402424101.

RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,

RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,

RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,

RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,

RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,

RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;

RT "Genome sequence of the enterobacterial phytopathogen Erwina

carotovora subsp. atroseptica and characterization of virulence

factors.";

RT Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).

DR EMBL: BX950851; CAG75345.1; -; Genomic DNA.

DR SMR: Q6D4B8; 2-207.

DR InterPro: IPR007451; DUF489.

```

DR Pfam: PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein
SQ SEQUENCE 212 AA; 22892 MW; A6781305F32B150 CRC64;

Query Match
Best Local Similarity 88.2%; Score 105; DB 2; Length 212;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 NYVDTTLAAGICGSRARLVQOLA 24
   |||:|||||:|||||:|||||:
Db 4 NYVEITTLAAGICGSRARLVQOLA 26

RESULT 13
Y1850_PASMU STANDARD; PRT; 203 AA.
ID Y1850_PASMU
AC Q9CJY8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DB Hypothetical UPF0274 protein PF1850.
GN OrderedLocustNames=PM1850;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
CX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE006223; AAK03934.1; -; Genomic_DNA.
DR HAMAP; MF 00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 203 AA; 22717 MW; 5EB91EBF0E181EDF CRC64;

Query Match
Best Local Similarity 86.6%; Score 103; DB 1; Length 203;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 NYVDTTLAAGICGSRARLVQOLA 24
   |||:|||||:|||||:|||||:
Db 3 NYVDTTLAAGICGSRARLVQOLA 25

RESULT 14
Y2805_PHOLL STANDARD; PRT; 208 AA.
ID Y2805_PHOLL
AC Q7N3B4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DB Hypothetical UPF0274 protein plu2805.
GN OrderedLocustNames=plu2805;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
CX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
DR STRAIN=TT01;

```

```

RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbr886;
RA Duchard E., Ruanlok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taouit S., Bocs S., Bouraux-Eude C., Chandier M., Charles J.-F.,
RA Daaba E., Deroose R., Derzelle S., Freysinet G., Gaudriault S.,
RA Medigue C., Ianolis A., Powell K., Siguler P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX571868; CAB15179.1; -; Genomic_DNA.
DR Photolysis; plu2805; -.
DR HAMAP; MF 00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 208 AA; 22831 MW; 3B484FA1FBFB36AD CRC64;

Query Match
Best Local Similarity 81.5%; Score 97; DB 1; Length 208;
Matches 19; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 NYVDTTLAAGICGSRARLVQOLA 24
   |||:|||||:|||||:|||||:
Db 4 NHVDTTLAAGICGSRARLVQOLA 26

RESULT 15
Y638_HAEIN STANDARD; PRT; 205 AA.
ID Y638_HAEIN
AC P44796;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DB Hypothetical UPF0274 protein HI0638.
GN OrderedLocustNames=HI0638;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
CX NCBI_TaxID=727;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kinsess E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd";
RL Science 269:496-512(1995).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U32747; AAC22298.1; -; Genomic_DNA.
DR PIR; I64155; I64155.

```

DR TIGR; H10638; -;
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 2323 MW; 81F31B0186BF82DA CRC64;

Query Match 79.8%; Score 95; DB 1; Length 205;
Best Local Similarity 78.3%; Pred. No. 8.4e-07;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 NYDITLALAGICQSAKLVQOLA 24
||:|||||:|||||
Db 3 NYHDIYTLALAGVQSAKLVHOLA 25

Search completed: January 23, 2006, 10:40:56
Job time : 76.5066 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:35:50 ; Search time 22.0087 Seconds
(without alignments)
90.156 Million cell updates/sec

Title: US-10-501-838A-25

Sequence: 1 NYVDITLALGICQSARLVQQLA 24

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5.COMB.pep.*
2: /cgn2_6/prodata/1/1aa/6.COMB.pep.*
3: /cgn2_6/prodata/1/1aa/H.COMB.pep.*
4: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep.*
5: /cgn2_6/prodata/1/1aa/RE.COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	92.4	217	2	US-09-489-039A-7789 Sequence 7789, Ap
2	88	73.9	216	2	US-09-543-681A-7993 Sequence 7993, Ap
3	56	47.1	138	2	US-09-270-767-58510 Sequence 58510, A
4	56	47.1	433	2	US-09-270-767-43172 Sequence 43172, A
5	48	40.3	359	1	US-07-937-609-19 Sequence 19, Appl
6	48	40.3	359	2	US-08-029-170-19 Sequence 19, Appl
7	48	40.3	359	2	US-09-443-745-19 Sequence 19, Appl
8	47	39.5	384	1	US-08-103-170-10 Sequence 10, Appl
9	47	39.5	387	1	US-08-196-989B-14 Sequence 14, Appl
10	47	39.5	387	2	US-08-760-936-14 Sequence 14, Appl
11	47	39.5	387	2	US-09-225-024-14 Sequence 14, Appl
12	47	39.5	387	2	US-09-826-509-527 Sequence 527, App
13	46	38.7	159	2	US-09-489-039A-10685 Sequence 10685, A
14	46	38.7	231	2	US-09-830-433A-32 Sequence 32, Appl
15	44	37.0	462	1	US-08-471-033-20 Sequence 20, Appl
16	44	37.0	462	1	US-08-471-044-20 Sequence 20, Appl
17	44	37.0	462	1	US-08-463-483A-20 Sequence 20, Appl
18	44	37.0	462	1	US-08-471-046A-20 Sequence 20, Appl
19	44	37.0	462	1	US-08-470-568B-20 Sequence 20, Appl
20	44	37.0	462	1	US-08-469-334-20 Sequence 20, Appl
21	44	37.0	462	2	US-09-300-529-20 Sequence 20, Appl
22	43	36.1	400	2	US-09-543-681A-4587 Sequence 4587, Ap
23	43	36.1	434	2	US-09-328-352-7817 Sequence 7817, Ap
24	42.5	35.7	427	2	US-09-248-796A-17203 Sequence 17203, A
25	42.5	35.7	762	2	US-09-228-986-114 Sequence 114, App
26	42.5	35.7	762	2	US-10-101-464A-114 Sequence 114, App
27	42	35.3	823	2	US-09-252-991A-23655 Sequence 23655, A

ALIGNMENTS

28	41.5	34.9	717	2	US-09-107-532A-6350 Sequence 6350, Ap
29	41	34.5	60	2	US-09-248-796A-37407 Sequence 27407, A
30	41	34.5	114	2	US-09-902-540-16640 Sequence 10640, A
31	41	34.5	231	2	US-09-830-433A-71 Sequence 71, Appl
32	41	34.5	309	2	US-09-902-540-12828 Sequence 12828, A
33	41	34.5	499	2	US-09-902-540-14780 Sequence 14780, A
34	41	34.5	588	2	US-09-605-703B-280 Sequence 280, App
35	41	34.5	588	2	US-09-605-703B-282 Sequence 282, App
36	41	34.5	741	2	US-09-252-991A-27062 Sequence 27062, A
37	40.5	34.0	236	2	US-09-583-110-3029 Sequence 3029, Ap
38	40.5	34.0	241	2	US-09-107-433-4815 Sequence 4815, Ap
39	40.5	34.0	742	2	US-09-215-694-12 Sequence 12, Appl
40	40.5	34.0	742	2	US-10-109-310-12 Sequence 12, Appl
41	40.5	34.0	742	2	US-09-487-558B-26 Sequence 26, Appl
42	40	33.6	191	2	US-09-489-039A-10634 Sequence 10634, A
43	40	33.6	207	2	US-09-248-796A-16052 Sequence 16052, A
44	40	33.6	262	2	US-09-813-453B-8 Sequence 8, Appl
45	40	33.6	303	2	US-09-543-681A-7924 Sequence 7924, Ap

RESULT 1
US-09-489-039A-7789
; Sequence 7789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004/001
; CURRENT FILING DATE: US/09/489, 039A
; PRIOR APPLICATION NUMBER: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7789
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match 92.4%; Score 110; DB 2; Length 217;
Best Local Similarity 91.3%; Pred. No. 2e+10; 0; Indels 0; Gaps 0;
Matches 21; Conservative 2; Mismatches 0;

QY 2 NYVDITLALGICQSARLVQQLA 24
DB 8 NYVDITLALGVCQARLVQQLA 30

RESULT 2
US-09-543-681A-7993
; Sequence 7993, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7993
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7993

Query Match 73.9%; Score 88; DB 2; Length 216;

Best Local Similarity 73.9%; Pred. No. 9.4e-07;
Matches 17; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NYDYDTLALAGICQASRLVQOLA 24
Db 11 DFRDITLALAGICQASRLVQOLA 33

RESULT 3

US-09-270-767-58510
; Sequence 58510, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58510
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-58510

Query Match 47.1%; Score 56; DB 2; Length 138;
Best Local Similarity 61.1%; Pred. No. 0.12;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYDYDTLALAGICQASRLVQOLA 18
Db 23 NYDYDTLALAGICQASRLVQOLA 40

RESULT 4

US-09-270-767-43172
; Sequence 43172, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43172
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43172

Query Match 47.1%; Score 56; DB 2; Length 433;
Best Local Similarity 61.1%; Pred. No. 0.47;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NYDYDTLALAGICQASRLVQOLA 18
Db 318 NYDYDTLALAGICQASRLVQOLA 335

RESULT 5

US-07-937-609-19
; Sequence 19, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937, 609
FILING DATE: 19920902

CLASSIFICATION: 436

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 40399/166 NIND
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: rat substance K receptor
US-07-937-609-19

Query Match 40.3%; Score 48; DB 1; Length 369;
Best Local Similarity 56.2%; Pred. No. 8.5;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NYDYDTLALAGICQASRLVQOLA 17
Db 69 NYDYDTLALAGICQASRLVQOLA 84

RESULT 6

US-08-029-170-19
; Sequence 19, Application US/08029170
; Patent No. 6169173
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIH
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-08-029-170-19

Query Match 40.3%; Score 48; DB 2; Length 369;
Best Local Similarity 56.2%; Pred. No. 8.5;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQA 17
DB 69 NYFIINLALADLCMAA 84

RESULT 7
US-09-443-745-19
Sequence 19, Application US/09443745
Patent No. 6706493
GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033

FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIH
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-09-443-745-19

Query Match 40.3%; Score 48; DB 2; Length 369;
Best Local Similarity 56.2%; Pred. No. 8.5;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQA 17
DB 69 NYFIINLALADLCMAA 84

RESULT 8
US-08-103-170-10
Sequence 10, Application US/08103170
Patent No. 5885824
GENERAL INFORMATION:
APPLICANT: Yamada, Tadataka
TITLE OF INVENTION: Recombinant Genomic Clones Encoding
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCLELLAND, WAIER & NEUSTADT,
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,170
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/633,060
FILING DATE: 24-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye, Jean-Paul
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-017-55
TELEPHONE: (703) 521-4500
TELEFAX: (703) 486-2347
TELEX: 24885 OPAT UR
INFORMATION FOR SEQ ID NO: 10:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Bos taurus
US-08-103-170-10

Query Match          39.5%; Score 47; DB 1; Length 384;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy      2 NYDITLALAGCOSA 17
Db      69 NYFIVNLADLCMAA 84

RESULT 9
US-08-196-989B-14
; Sequence 14, Application US/08196989B
; Patent No. 5585476
; GENERAL INFORMATION:
; APPLICANT: MacLennan, A. John
; TITLE OF INVENTION: Molecular Cloning and Expression of
; TITLE OF INVENTION: G-Protein Coupled Receptors
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,989B
; FILING DATE: 15-FEB-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Lloyd, Jeff
; REGISTRATION NUMBER: 35,589
; REFERENCE/DOCKET NUMBER: MAC-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-196-989B-14

Query Match          39.5%; Score 47; DB 1; Length 387;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy      2 NYDITLALAGCOSA 17
Db      69 NYFIVNLADLCMAA 84

RESULT 10
US-08-760-936-14
; Sequence 14, Application US/08760936
; Patent No. 5856443
; GENERAL INFORMATION:
```

```

; APPLICANT: MacLennan, A. John
; TITLE OF INVENTION: Molecular Cloning and Expression of
; TITLE OF INVENTION: G-Protein Coupled Receptors
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,936
; FILING DATE: December 6, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Doran R.
; REGISTRATION NUMBER: 38,261
; REFERENCE/DOCKET NUMBER: MAC-100C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-760-936-14

Query Match          39.5%; Score 47; DB 1; Length 387;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy      2 NYDITLALAGCOSA 17
Db      69 NYFIVNLADLCMAA 84

RESULT 11
US-09-225-024-14
; Sequence 14, Application US/09225024
; Patent No. 6518414
; GENERAL INFORMATION:
; APPLICANT: MacLennan, A. John
; TITLE OF INVENTION: Molecular Cloning and Expression of
; TITLE OF INVENTION: G-Protein Coupled Receptors
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/225,024
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/760,936
; FILING DATE: 6-DEC-1996
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/196,989
FILING DATE: 15-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: MAC-100C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 352-375-8100
TELEFAX: 352-372-5800
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-225-024-14

Query Match 39.5%; Score 47; DB 2; Length 387;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQSA 17
DB 69 NYFVNLALADLCMAA 84

RESULT 12
US-09-826-509-527
Sequence 527, Application US/09826509
Patent No. 6806054
GENERAL INFORMATION:
APPLICANT: Lehmann-Brunsmann, Karin
APPLICANT: Lin, I-Lin
TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
FILE REFERENCE: AREN-207
CURRENT APPLICATION NUMBER: US/09/826,509
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: 60/195,747
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 589
SOFTWARE: Patent Version 2.1
SEQ ID NO 527
LENGTH: 398
TYPE: PRP
ORGANISM: Homo sapiens
US-09-826-509-527

Query Match 39.5%; Score 47; DB 2; Length 398;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQSA 17
DB 69 NYFVNLALADLCMAA 84

RESULT 13
US-09-489-039A-10685
Sequence 10685, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10685
LENGTH: 169
TYPE: PRP
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10685

Query Match 38.7%; Score 46; DB 2; Length 169;
Best Local Similarity 71.4%; Pred. No. 7.4;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQ 15
DB 70 NHRITLALAGFRQ 83

RESULT 14
US-09-830-433A-32
Sequence 32, Application US/09830433A
Patent No. 6835384
GENERAL INFORMATION:
APPLICANT: AUME et al.
TITLE OF INVENTION: Nucleic acids and polypeptides specific for pathogenic
FILE REFERENCE: P07180US00/BAS
CURRENT APPLICATION NUMBER: US/09/830,433A
CURRENT FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: FR 98 13 693
PRIOR FILING DATE: 1998-10-30
NUMBER OF SEQ ID NOS: 129
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 32
LENGTH: 231
TYPE: PRP
ORGANISM: Neisseria meningitidis
US-09-830-433A-32

Query Match 38.7%; Score 46; DB 2; Length 231;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YDITLALAGICQAR 18
DB 195 YDILASAPAPICRQAR 210

RESULT 15
US-08-471-033-20
Sequence 20, Application US/08471033
Patent No. 5770696
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Koziele, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Batruch, Juan J
TITLE OF INVENTION: No. 5770696el Peptidic Proteins and Strains
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSER: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,033
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: P-40,403
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLV3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-471-033-20

Query Match 37.0%; Score 44; DB 1; Length 462;
Best Local Similarity 50.0%; Pred. NO. 52;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 2 NYDYITLALAGICQ 15
||:||||:|
Db 107 NYKEITPSMAGSCE 120

Search completed: January 23, 2006, 10:46:49
Job time : 22.0087 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:53:39 ; Search time 3.35371 Seconds
(without alignments)
72.521 Million cell updates/sec

Title: US-10-501-838a-25

Perfect score: 119

Sequence: 1 MWYDITLALGICQSARLVQQLA 24

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/pcodata/2/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/pcodata/2/pubppaa/US06_NEW_PUB.pep.*
3: /cgn2_6/pcodata/2/pubppaa/US07_NEW_PUB.pep.*
4: /cgn2_6/pcodata/2/pubppaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/pcodata/2/pubppaa/US09_NEW_PUB.pep.*
6: /cgn2_6/pcodata/2/pubppaa/US10_NEW_PUB.pep.*
7: /cgn2_6/pcodata/2/pubppaa/US11_NEW_PUB.pep.*
8: /cgn2_6/pcodata/2/pubppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	36.1	42	US-10-512-295A-3	Sequence 3, Appl1
2	43	36.1	185	US-10-453-372-896	Sequence 896, App
3	43	36.1	194	US-10-453-372-894	Sequence 894, App
4	43	36.1	218	US-10-453-372-892	Sequence 892, App
5	41	34.5	305	US-10-520-820-1	Sequence 1, Appl1
6	40.5	34.0	365	US-11-000-463-241	Sequence 241, App
7	40.5	34.0	365	US-11-000-463-713	Sequence 713, App
8	40	33.6	63	US-10-467-657-6638	Sequence 6638, App
9	39	32.8	1015	US-10-957-569-51	Sequence 51, Appl1
10	38.5	32.4	447	US-10-858-730-219	Sequence 219, App
11	38	31.9	203	US-10-453-372-890	Sequence 890, App
12	38	31.9	655	US-10-055-877-40	Sequence 40, Appl1
13	38	31.9	678	US-10-055-877-202	Sequence 202, App
14	38	31.9	687	US-10-055-877-199	Sequence 199, App
15	38	31.9	687	US-10-055-877-201	Sequence 201, App
16	38	31.9	787	US-10-467-657-2832	Sequence 2832, App
17	37	31.1	367	US-10-821-334-1569	Sequence 1569, App
18	36.5	30.7	447	US-11-055-822-94	Sequence 94, Appl1
19	36.5	30.7	481	US-10-467-657-4660	Sequence 4660, App
20	36.5	30.7	522	US-11-055-822-1100	Sequence 1100, App
21	36	30.3	228	US-10-467-657-1242	Sequence 1242, App
22	36	30.3	231	US-11-082-389-242	Sequence 242, App
23	36	30.3	301	US-10-467-657-7356	Sequence 7356, App
24	36	30.3	350	US-10-467-657-3256	Sequence 3256, App
25	36	30.3	434	US-10-467-657-6894	Sequence 6894, App

26	36	30.3	473	7	US-11-069-642-5	Sequence 5, Appl1
27	36	30.3	481	6	US-10-763-712A-4	Sequence 4, Appl1
28	36	30.3	883	6	US-10-793-626-754	Sequence 754, App
29	36	30.3	919	6	US-10-858-730-206	Sequence 206, App
30	36	30.3	1268	7	US-11-052-554A-1	Sequence 1, Appl1
31	35.5	29.8	341	6	US-10-467-657-154	Sequence 154, App
32	35.5	29.8	341	6	US-10-467-657-5822	Sequence 5822, App
33	35.5	29.8	341	6	US-10-467-657-7476	Sequence 7476, App
34	35.5	29.8	546	6	US-10-821-234-902	Sequence 902, App
35	35.5	29.8	884	6	US-10-878-536A-9	Sequence 9, Appl1
36	35.5	29.8	974	6	US-10-821-234-1152	Sequence 1152, App
37	35.5	29.8	1259	6	US-10-467-657-5510	Sequence 5510, App
38	35	29.4	250	6	US-10-793-626-2310	Sequence 2310, App
39	35	29.4	269	7	US-11-151-482-5	Sequence 5, Appl1
40	35	29.4	270	6	US-10-467-657-7100	Sequence 7100, App
41	35	29.4	293	6	US-10-967-671-15	Sequence 15, Appl1
42	35	29.4	305	7	US-11-000-463-391	Sequence 391, App
43	35	29.4	305	7	US-11-000-463-863	Sequence 863, App
44	35	29.4	332	6	US-10-454-437-308	Sequence 308, App
45	35	29.4	356	6	US-10-793-626-2054	Sequence 2054, App

ALIGNMENTS

```

RESULT 1
US-10-512-295A-3
; Sequence 3, Application US/10512295A
; Publication No. US20050245727A1
; GENERAL INFORMATION:
; APPLICANT: Baltzer, Lars
; APPLICANT: Dolphin, Gunnar
; APPLICANT: Liedberg, Bo
; APPLICANT: Lundstrom, Ingemar
; TITLE OF INVENTION: NOVEL POLYPEPTIDE SCAFFOLDS AND USE THEREOF
; FILE REFERENCE: 5848.18USMO
; CURRENT APPLICATION NUMBER: US/10/512,295A
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/SE03/00507
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: SE 0200968-6
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide XE3
; US-10-512-295A-3

Query Match
Best Local Similarity 52.9%; Pred. No. 0.68;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      8 LALAGICQSARLVQQLA 24
      |||||:|:|:|
Db      16 LAARGPCDAQLAEQLA 32

RESULT 2
US-10-453-372-896
; Sequence 896, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
; APPLICANT: Alcobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23

```

```

; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Curaseqblast version 0.1
; SEQ ID NO: 896
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-896
```

```

Query Match          36.1%; Score 43; DB 6; Length 185;
Best Local Similarity 52.4%; Pred. No. 3.3;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```

Qy      4 YDITLALGICQSARLVQOLA 24
Db      67 YDSLALDGHIGSARALMVVA 87
```

```

RESULT 3
US-10-453-372-894
; Sequence 894, Application US/10453372
; Publication No. US2006003323A1
; GENERAL INFORMATION:
; APPLICANT: Alcobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Curaseqblast version 0.1
; SEQ ID NO: 894
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

US-10-453-372-894
```

```

Query Match          36.1%; Score 43; DB 6; Length 194;
Best Local Similarity 52.4%; Pred. No. 3.5;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```

Qy      4 YDITLALGICQSARLVQOLA 24
Db      47 YDSLALDGHIGSARALMVVA 67
```

```

RESULT 4
US-10-453-372-892
; Sequence 892, Application US/10453372
; Publication No. US2006003323A1
; GENERAL INFORMATION:
; APPLICANT: Alcobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Curaseqblast version 0.1
; SEQ ID NO: 892
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-892
```

```

Query Match          36.1%; Score 43; DB 6; Length 218;
Best Local Similarity 52.4%; Pred. No. 3.9;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```

Qy      4 YDITLALGICQSARLVQOLA 24
Db      71 YDSLALDGHIGSARALMVVA 91
```

```

RESULT 5
US-10-520-820-1
; Sequence 1, Application US/10520820
; Publication No. US2006003393A1
; GENERAL INFORMATION:
; APPLICANT: MUTHAIBILIS S.A.
; TITLE OF INVENTION: Pathogenicity determinants which can be used as targets for devel
; TITLE OF INVENTION: means for preventing and controlling bacterial infections and/or
; FILE REFERENCE: 1621
; CURRENT APPLICATION NUMBER: US/10/520,820
; PRIOR FILING DATE: 2005-01-07
; PRIOR APPLICATION NUMBER: PCT/EP2003/008209
; PRIOR FILING DATE: 2003-07-09
```

/ PRIOR APPLICATION NUMBER: FR 0208636
 / PRIOR FILING DATE: 2002-07-09
 / NUMBER OF SEQ ID NOS: 32
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO 1
 / LENGTH: 305
 / TYPE: PRT
 / ORGANISM: Escherichia coli
 US-10-520-820-1

Query Match 34.5%; Score 41; DB 6; Length 305;
 Best Local Similarity 42.9%; Pred. No. 12;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 4 YDITLALGICQSARLVQOLA 24
 DB 69 YDAVIDAGSLVKSALVTRLA 89

RESULT 6
 US-11-000-463-241
 / Sequence 241, Application US/11000463
 / Publication No. US20050266423A1
 / GENERAL INFORMATION:
 / APPLICANT: Tang, Y Tom
 / APPLICANT: Liu, Chenghua
 / APPLICANT: Asundi, Vinod
 / APPLICANT: Chen, Rui-hong
 / APPLICANT: Qian, Xiaohong B.
 / APPLICANT: Wang, Zhiwei
 / APPLICANT: Wehrman, Tom
 / APPLICANT: Zhang, Jie
 / APPLICANT: Zhou, Ping
 / APPLICANT: Cao, Yi-Cheng
 / APPLICANT: Drmanac, Radoje T.
 / TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
 / FILE REFERENCE: 785CIP4CN
 / CURRENT APPLICATION NUMBER: US/11/000,463
 / PRIOR FILING DATE: 2004-11-29
 / PRIOR APPLICATION NUMBER: 10/291,265
 / PRIOR FILING DATE: 2002-11-08
 / PRIOR APPLICATION NUMBER: PCT/US01/02623
 / PRIOR FILING DATE: 2001-01-25
 / PRIOR APPLICATION NUMBER: 09/922,279
 / PRIOR FILING DATE: 2001-08-03
 / PRIOR APPLICATION NUMBER: 09/491,404
 / PRIOR FILING DATE: 2000-01-25
 / PRIOR APPLICATION NUMBER: 09/617,746
 / PRIOR FILING DATE: 2000-07-17
 / PRIOR APPLICATION NUMBER: 09/631,451
 / PRIOR FILING DATE: 2000-08-03
 / PRIOR APPLICATION NUMBER: 09/633,870
 / PRIOR FILING DATE: 2000-09-15
 / NUMBER OF SEQ ID NOS: 944
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 241
 / LENGTH: 365
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-11-000-463-241

Query Match 34.0%; Score 40.5; DB 7; Length 365;
 Best Local Similarity 43.5%; Pred. No. 18;
 Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 5 DITLALGIC--QSARLVQOLA 24
 DB 67 DLMGLAGFCGSMASGHLFKOMA 89

RESULT 7
 US-11-000-463-713
 / Sequence 713, Application US/11000463

/ Publication No. US20050266423A1
 / GENERAL INFORMATION:
 / APPLICANT: Tang, Y Tom
 / APPLICANT: Liu, Chenghua
 / APPLICANT: Asundi, Vinod
 / APPLICANT: Chen, Rui-hong
 / APPLICANT: Qian, Xiaohong B.
 / APPLICANT: Wang, Zhiwei
 / APPLICANT: Wehrman, Tom
 / APPLICANT: Zhang, Jie
 / APPLICANT: Zhou, Ping
 / APPLICANT: Cao, Yi-Cheng
 / APPLICANT: Drmanac, Radoje T.
 / TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
 / FILE REFERENCE: 785CIP4CN
 / CURRENT APPLICATION NUMBER: US/11/000,463
 / PRIOR FILING DATE: 2004-11-29
 / PRIOR APPLICATION NUMBER: 10/291,265
 / PRIOR FILING DATE: 2002-11-08
 / PRIOR APPLICATION NUMBER: PCT/US01/02623
 / PRIOR FILING DATE: 2001-01-25
 / PRIOR APPLICATION NUMBER: 09/922,279
 / PRIOR FILING DATE: 2001-08-03
 / PRIOR APPLICATION NUMBER: 09/491,404
 / PRIOR FILING DATE: 2000-01-25
 / PRIOR APPLICATION NUMBER: 09/617,746
 / PRIOR FILING DATE: 2000-07-17
 / PRIOR APPLICATION NUMBER: 09/631,451
 / PRIOR FILING DATE: 2000-08-03
 / PRIOR APPLICATION NUMBER: 09/633,870
 / PRIOR FILING DATE: 2000-09-15
 / NUMBER OF SEQ ID NOS: 944
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 713
 / LENGTH: 365
 / TYPE: PRT
 / ORGANISM: Homo sapiens
 US-11-000-463-713

Query Match 34.0%; Score 40.5; DB 7; Length 365;
 Best Local Similarity 43.5%; Pred. No. 18;
 Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 5 DITLALGIC--QSARLVQOLA 24
 DB 67 DLMGLAGFCGSMASGHLFKOMA 89

RESULT 8
 US-10-467-657-6638
 / Sequence 6638, Application US/10467657
 / Publication No. US20050260581A1
 / GENERAL INFORMATION:
 / APPLICANT: CHIRON SPA
 / APPLICANT: FONTANA Maria Rita
 / APPLICANT: PIZZA Mariagrazia
 / APPLICANT: MASIGNANI Vega
 / APPLICANT: MONACI Elisabetta
 / TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 / FILE REFERENCE:
 / CURRENT APPLICATION NUMBER: US/10/467,657
 / PRIOR FILING DATE: 2003-08-11
 / PRIOR APPLICATION NUMBER: GB-0103424.8
 / PRIOR FILING DATE: 2001-02-12
 / NUMBER OF SEQ ID NOS: 9218
 / SOFTWARE: SeqWin99, version 1.04
 / SEQ ID NO 6638
 / LENGTH: 63
 / TYPE: PRT
 / ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-6638

Query Match 33.6%; Score 40; DB 6; Length 63;

Best Local Similarity 53.3%; Pred. No. 3.3;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YDITLALAGICQSA 17
|||:|:|:|:
Db 46 YNITVASTVKCQTA 60

RESULT 9

US-10-957-569-51
; Sequence 51, Application US/10957569
; Publication No. US20050246785A1
; GENERAL INFORMATION:
; APPLICANT: COOK, Zhihong et al.
; TITLE OF INVENTION: PROMOTER, PROMOTER CONTROL ELEMENTS, AND COMBINATIONS, AND USES
; FILE REFERENCE: 2750-1577PUS3
; CURRENT APPLICATION NUMBER: US/10/957,569
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 10/950,321
; PRIOR FILING DATE: 2004-09-23
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-957-569-51

Query Match 32.8%; Score 39; DB 6; Length 1015;
Best Local Similarity 35.0%; Pred. No. 92;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 5 DITLALGICQSA RLVQQLA 24
|||:|:|:|:|:|:
Db 172 DMTLMIDGVCARSLIVGIA 191

RESULT 10

US-10-858-730-219
; Sequence 219, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 219
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-858-730-219

Query Match 32.4%; Score 38.5; DB 6; Length 447;
Best Local Similarity 33.3%; Pred. No. 47;
Matches 10; Conservative 5; Mismatches 4; Indels 11; Gaps 1;

Qy 2 NYVDITL-----ALAGICQSA RLV 20
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 9 NYDMLKRNAGPEPFHQAVALDESLKTV 38

RESULT 11

US-10-453-372-890
; Sequence 890, Application US/10453372
; Publication No. US2006003323A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Curaseqdist version 0.1
; SEQ ID NO 890
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-890

Query Match 31.9%; Score 38; DB 6; Length 203;
Best Local Similarity 47.6%; Pred. No. 24;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 4 YDITLALGICQSA RLVQQLA 24
|||:|:|:|:|:|:
Db 67 YDSLALDGRQARALMVVA 87

RESULT 12

US-10-055-877-40
; Sequence 40, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tcherev, Vellizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zethusen, Bryan
; APPLICANT: Andrew, David


```
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eissen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Boldog, Ferenc
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 655
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-877-40

Query Match          31.9%; Score 38; DB 6; Length 655;
Best Local Similarity 64.3%; Pred. No. 84;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      10 LAGICQARIVQQL 23
DB      558 LVGIQRAQLVQAL 571

RESULT 13
US-10-055-877-202
Sequence 202, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchervnev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
```

```
APPLICANT: Guo, Xiaojia
APPLICANT: Zernusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eissen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Boldog, Ferenc
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 202
LENGTH: 678
TYPE: PRT
ORGANISM: Oryctolagus cuniculus
US-10-055-877-202

Query Match          31.9%; Score 38; DB 6; Length 678;
Best Local Similarity 64.3%; Pred. No. 88;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      10 LAGICQARIVQQL 23
DB      590 LVGIQRAQLVQAL 603

RESULT 14
US-10-055-877-199
Sequence 199, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchervnev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
```

APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zernusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Etsen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 199
LENGTH: 687
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-877-199

Query Match 31.9%; Score 38; DB 6; Length 687;
Best Local Similarity 64.3%; Pred. No. 89;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 10 LAGICQASRLVQQL 23
Db 590 LVGIQRAQLVQAL 603

RESULT 15
US-10-055-877-201
Sequence 201, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei

APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zernusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Etsen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 201
LENGTH: 687
TYPE: PRT
ORGANISM: Oryctolagus cuniculus
US-10-055-877-201

Query Match 31.9%; Score 38; DB 6; Length 687;
Best Local Similarity 64.3%; Pred. No. 89;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 10 LAGICQASRLVQQL 23
Db 590 LVGIQRAQLVQAL 603

Search completed: January 23, 2006, 12:09:42
Job time : 3.35371 secs

CC stimulating factor (GM-CSF), enkephalin, dargirin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC cyostatic or neurotropic activities, cardiant, antiarteriosclerotic, osteopathic,
 CC Becherichia coli and is penetrating peptide 30 of the invention.
 XX
 SQ Sequence 24 AA;
 Query Match 100.0%; Score 119; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 2.5e-12;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MNYDYDTLALAGICQSARLVQOLA 24
 Db 1 MNYDYDTLALAGICQSARLVQOLA 24
 RESULT 2
 AEB08242
 ID AEB08242 standard; peptide; 24 AA.
 AC AEB08242;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DE Becherichia coli YCFC penetrating peptide 30, SEQ ID NO: 25.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease; osteoporosis;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiant; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antipneumatic; cyostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Becherichia coli.
 XX
 PN US2005136103-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 16-SEP-2004; 2004US-00942300.
 XX
 PR 17-SEP-2003; 2003US-00664989.
 XX
 PR 17-SEP-2003; 2003US-00665184.
 XX
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 DR WPI; 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 25; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Becherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 SQ Sequence 24 AA;
 Query Match 100.0%; Score 119; DB 9; Length 24;
 Best Local Similarity 100.0%; Pred. No. 2.5e-12;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MNYDYDTLALAGICQSARLVQOLA 24
 Db 1 MNYDYDTLALAGICQSARLVQOLA 24
 RESULT 3
 ADB16918
 ID ADB16918 standard; peptide; 31 AA.
 XX
 AC ADB16918;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Becherichia coli peptide 3 coupled to imaging compound linker, IBW-002V1.
 XX
 KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiant; antiarteriosclerotic; osteopathic; cyostatic; neurotropic;
 KW imaging linker; penetrating peptide; IBW-002V1.
 XX
 OS Synthetic.
 XX
 PN Becherichia coli.
 XX
 PD 14-AUG-2003.
 XX
 PF 07-FEB-2003; 2003WO-IB000968.
 XX
 PR 07-FEB-2002; 2002US-0355396P.
 XX
 PT

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 XX WPI, 2003-697452/66.
 DR New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX Example 3; Page 40; 60pp; English.
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/macocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antihypertensive, osteopathic,
 CC cytotactic or neurotropic activities. This peptide sequence is IBM-002V1,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 XX Sequence 31 AA:
 SO
 Query Match 100.0%; Score 119; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 3.3e-12;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNYDITLALAGICQSARLVQQLA 24
 DB 1 MNYDITLALAGICQSARLVQQLA 24
 RESULT 4
 AEB08252
 ID AEB08252 standard; peptide; 31 AA.
 AC AEB08252;
 XX
 DT 25-ANG-2005 (first entry)
 XX Penetrating peptide SEQ: 35 used in composition for mucosal vaccination.
 DE
 XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degeneration; musculoskeletal disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimers disease;
 XX neurotropic; neurological disease; parkinsons disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; antihypertensive;
 XX coronary artery disease; cardiac; vasodilator; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotoxic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; virolytic;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antirheumatic; cytostatic;
 XX antiinflammatory; hepatotropic; hepatitis B virus infection.

XX Unidentified.
 OS
 XX Key Location/Qualifiers
 XX Misc-difference 1
 FT Misc-difference 1
 FT /note= "N-terminal acylated"
 FT /note= "Optionally C-terminal amide, optionally the free
 FT amino group of lysine is acylated with a fatty acid"
 XX US2005136103-A1.
 XX 23-JUN-2005.
 XX 16-SEP-2004; 2004US-00942300.
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX (BENS/) BEN-SASSON S A.
 XX (COHR/) COHEN E.
 XX Ben-Sasson SA, Cohen E;
 XX WPI, 2005-444089/45.
 DR Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX Claim 63; SEQ ID NO 35; 59pp; English.
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimers disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is a penetrating peptide (IBM-002V1)
 CC used in the composition for mucosal vaccination using a counter anion and
 CC a penetrating peptide.
 XX Sequence 31 AA:
 SO
 Query Match 100.0%; Score 119; DB 9; Length 31;
 Best Local Similarity 100.0%; Pred. No. 3.3e-12;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNYDITLALAGICQSARLVQQLA 24
 DB 1 MNYDITLALAGICQSARLVQQLA 24
 RESULT 5
 ADB16888
 ID ADB16888 standard; peptide; 23 AA.
 XX ADB16888;
 AC

XX 20-NOV-2003 (first entry)
 XX Escherichia coli YCFC penetrating peptide 3.
 DE
 XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 KW infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; arteriosclerotic;
 KW osteopathic; cytosolic; motropic.
 XX Escherichia coli.
 OS
 XX WO2003066859-A2.
 PN
 XX 14-AUG-2003.
 PD
 XX 07-FEB-2003; 2003WO-1B00968.
 PP
 XX 07-FEB-2003; 2002US-0355396P.
 PR
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA
 XX Ben-Sasson SA, Cohen E;
 XX MPI, 2003-697452/66.
 DR
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PT
 PS Claim 2; Page 14; 60pp; English.
 PS
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC antitoxigenic, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dargargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, arteriosclerotic, osteopathic,
 CC cytosolic or motropic activities. This peptide is from YCFC of
 CC Escherichia coli and is penetrating peptide 3 of the invention.
 XX
 XX Sequence 23 AA:
 SQ
 Query Match 95.8%; Score 114; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1,66-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW motropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; arteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; anitaneimic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytosolic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX Escherichia coli.
 OS
 XX US2005136103-A1.
 PN
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 PF
 XX 17-SEP-2003; 2003US-00664989.
 PR
 XX 17-SEP-2003; 2003US-00665184.
 PR
 XX 17-SEP-2003; 2003US-0503615P.
 PA
 XX (BENS/) BEN-SASSON S A.
 PA
 XX (COHE/) COHEN E.
 PI
 XX Ben-Sasson SA, Cohen E;
 XX MPI, 2005-444089/45.
 DR
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 PT
 XX Claim 53; SEQ ID NO 3; 59pp; English.
 PS
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorder, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 XX Sequence 23 AA;
 SQ

Query Match 95.8%; Score 114; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.6e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDITTLALAGICOSARLVQOLA 24
 |||||
 DB 1 NYDITTLALAGICOSARLVQOLA 23

RESULT 7

ADBI6922
 ID ADBI6922 standard; peptide; 25 AA.

AC ADBI6922;

DT 20-NOV-2003 (first entry)

DE Escherichia coli peptide 3 linked to recombinant human insulin.

XX recombinant human insulin; epithelial; endothelial; tight junction;
 XX diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 XX cardiovascular; haematological; endocrine disorder; obesity;
 XX neoplastic disease; neuroprotective; cardiatic; antiarteriosclerotic;
 XX osteopathic; cytostatic; nootropic; penetrating peptide; diabetes.

OS Synthetic.
 XX Escherichia coli.

FT Key Location/Qualifiers
 FT Peptide 1..23
 FT Modified-site 25 /note= "Penetrating peptide 3"

FT /note= "Recombinant human insulin peptide coupled to the
 C-terminus"

PN WO2003066859-A2.

PD 14-AUG-2003.

PP 07-FEB-2003; 2003WO-IB000968.

PR 07-FEB-2002; 2002US-0355396P.

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PI Ben-Sasson SA, Cohen E;

DR WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.

PS Example 8; Page 42; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, haematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX Accordingly, the peptides of this invention can be used in compositions
 XX that have neuroprotective, cardiatic, antiarteriosclerotic, osteopathic,
 XX cytostatic or nootropic activities. This peptide sequence consists of the
 XX Escherichia coli penetrating peptide 3 coupled to recombinant human

CC insulin in the absence of a detachable linker peptide as a cleavage site,
 CC used to deliver insulin across mucosal epithelia as a treatment for
 CC diabetes, in an exemplification of the invention.

CC Sequence 25 AA;

Query Match 95.8%; Score 114; DB 6; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.8e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDITTLALAGICOSARLVQOLA 24
 |||||
 DB 1 NYDITTLALAGICOSARLVQOLA 23

RESULT 8

ADBI6906
 ID ADBI6906 standard; peptide; 25 AA.

AC ADBI6906;

DT 20-NOV-2003 (first entry)

DE Escherichia coli peptide 3 linked to linearised insulin receptor peptide.

XX linearised insulin receptor; epithelial; endothelial; tight junction;
 XX diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 XX cardiovascular; haematological; endocrine disorder; obesity;
 XX neoplastic disease; neuroprotective; cardiatic; antiarteriosclerotic;
 XX osteopathic; cytostatic; nootropic; penetrating peptide; diabetes.

OS Synthetic.
 XX Escherichia coli.

FT Key Location/Qualifiers
 FT Peptide 1..23
 FT Modified-site 25 /note= "Linearised insulin receptor peptide coupled to
 the C-terminus"

FT /note= "Linearised insulin receptor peptide coupled to
 the C-terminus"

PN WO2003066859-A2.

PD 14-AUG-2003.

PP 07-FEB-2003; 2003WO-IB000968.

PR 07-FEB-2002; 2002US-0355396P.

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PI Ben-Sasson SA, Cohen E;

DR WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.

PS Example 9; Page 43; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin

CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
CC cyrostatic or neurotropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to a linearised insulin
CC receptor, used to deliver insulin across mucosal epithelia as a treatment
CC for diabetes, in an exemplification of the invention.
XX
SQ Sequence 25 AA;
Query Match 95.8%; Score 114; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 NYDYDTTLAAGICQSARLVQOLA 24
Db 1 NYDYDTTLAAGICQSARLVQOLA 23
RESULT 9
ADBI6904
ID ADBI6904 standard; peptide; 26 AA.
XX
AC ADBI6904;
XX
DT 20-NOV-2003 (first entry)
XX
DE E_coli penetrating peptide 3 linked to heparin without a cleavage site.
XX
XX heparin; epithelial; endothelial; tight junction; diabetes; infertility;
XX hormone; vitamin deficiency; neurodegenerative; cardiovascular;
XX haematological; endocrine disorder; obesity; neoplastic disease;
XX neuroprotective; cardiant; antiarteriosclerotic; osteopathic; cyrostatic;
XX neurotropic; penetrating peptide.
OS
OS Synthetic.
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /note= "Penetrating peptide 3"
FT Modified-site 26
FT /note= "Heparin peptide coupled to the C-terminus"
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Saason SA, Cohen E;
XX
XX WPI; 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Example 10; Page 44; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells
XX sealed by tight junctions. This peptide is derived from a bacterial
XX toxin, an integral membrane or extracellular protein and can comprise an
XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX or enzyme. The effector molecule, however, can comprise for example

CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
CC cyrostatic or neurotropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to heparin in the absence
CC of a detachable linker peptide as a cleavage site, used to deliver low
CC molecular weight heparin across mucosal epithelia in an exemplification
CC of the invention.
XX
SQ Sequence 26 AA;
Query Match 95.8%; Score 114; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 NYDYDTTLAAGICQSARLVQOLA 24
Db 1 NYDYDTTLAAGICQSARLVQOLA 23
RESULT 10
ADBI6921
ID ADBI6921 standard; peptide; 29 AA.
XX
AC ADBI6921;
XX
DT 20-NOV-2003 (first entry)
XX
XX E_coli peptide 3 linked to recombinant human insulin via cleavage site.
XX
XX recombinant human insulin; epithelial; endothelial; tight junction;
XX diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
XX cardiovascular; haematological; endocrine disorder; obesity;
XX neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
XX osteopathic; cyrostatic; neurotropic; detachable linker;
XX penetrating peptide; diabetes.
OS
OS Synthetic.
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /note= "Penetrating peptide 3"
FT Cleavage-site 26..29
FT /note= "Detachable linker peptide"
FT Modified-site 29
FT /note= "Recombinant human insulin peptide coupled to the
FT C-terminus"
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Saason SA, Cohen E;
XX
XX WPI; 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Example 8; Page 42; 60pp; English.
XX

CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticosugulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dargargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, hematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide sequence consists of the
 CC Escherichia coli penetrating peptide 3 coupled to recombinant human
 CC insulin via the detachable linker peptide as a cleavage site, used to
 CC deliver insulin across mucosal epithelia as a treatment for diabetes, in
 CC an exemplification of the invention.

XX Sequence 29 AA;

Query Match 95.8%; Score 114; DB 6; Length 29;

Best Local Similarity 100.0%; Pred. No. 2.1e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDDITLALAGICQSARLVQQLA 24

DB 1 NYDDITLALAGICQSARLVQQLA 23

RESULT 11

ADBI6903 ADBI6903 standard; peptide; 30 AA.

AC ADBI6903;

DT 20-NOV-2003 (first entry)

DE E.coli penetrating peptide 3 linked to heparin through a cleavage site.

XX heparin; epithelial; endothelial; tight junction; diabetes; infertility;

KW hormone; vitamin deficiency; neurodegenerative; cardiovascular;

KW haematological; endocrine disorder; obesity; neoplastic disease;

KW neuroprotective; cardiant; antiarteriosclerotic; osteopathic;

KW neurotropic; detachable linker; penetrating peptide.

OS Synthetic.

OS Escherichia coli.

XX Key

XX Peptide

XX Location/Qualifiers

XX 1..23

XX /note= "Penetrating peptide 3"

XX Cleavage-site

XX 26..29

XX /note= "detachable linker peptide"

XX Modified-site

XX 30

XX /note= "Heparin peptide coupled to the C-terminus"

XX MO2003066859-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003MO-IB000968.

XX 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Ben-Sasson SA, Cohen E;

XX PI

XX XX

DR MPI; 2003-697452/66.

XX New penetrating peptide, useful for preparing a composition for treating

PT or preventing e.g. endocrine disorders.

XX Example 10; Page 44; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticosugulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dargargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, hematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide sequence consists of the
 CC Escherichia coli penetrating peptide 3 coupled to heparin via the
 CC detachable linker peptide as a cleavage site, used to deliver low
 CC molecular weight heparin across mucosal epithelia in an exemplification
 CC of the invention.

XX Sequence 30 AA;

Query Match 95.8%; Score 114; DB 6; Length 30;

Best Local Similarity 100.0%; Pred. No. 2.2e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDDITLALAGICQSARLVQQLA 24

DB 1 NYDDITLALAGICQSARLVQQLA 23

RESULT 12

ADBI6905 ADBI6905 standard; peptide; 30 AA.

AC ADBI6905;

DT 20-NOV-2003 (first entry)

DE Escherichia coli peptide 3 coupled to imaging compound linker, IBM-002.

XX epithelial; endothelial; tight junction; diabetes; infertility; hormone;

KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;

KW endocrine disorder; obesity; neoplastic disease; neuroprotective;

KW cardiant; antiarteriosclerotic; osteopathic; neurotropic;

KW imaging linker; penetrating peptide; IBM-002.

OS Synthetic.

OS Escherichia coli.

XX Key

XX Modified-site

XX 1

XX /note= "N-terminal acetyl"

XX Modified-site

XX 30

XX /note= "C-terminal amide"

XX MO2003066859-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003MO-IB000968.

XX 07-FEB-2002; 2002US-0355396P.

XX PR

XX XX

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2003-697452/66.
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX Example 1; Page 33; 60pp; English.
 XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide sequence is IBW-002,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 XX Sequence 30 AA;

Query Match 95.8%; Score 114; DB 6; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDYITLALAGICQSAARLVQOLA 24
 Db 1 NYDYITLALAGICQSAARLVQOLA 23

RESULT 13

AEB08239 standard; peptide; 30 AA.

XX AEB08239;
 XX 25-AUG-2005 (first entry)
 XX Penetrating peptide SEQ: 22 used in composition for mucosal vaccination.
 DE
 XX Pharmacological; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration, musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotonic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; vironcide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cystostatic;

KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX Unidentified.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT /note= "N-terminal acylated"
 FT Misc-difference 30
 FT /note= "Optionally C-terminal amide, optionally the free
 FT amino group of lysine is acylated with a fatty acid"
 FT
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 DR WPI; 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 63; SEQ ID NO 22; 59pp; English.

XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC pathological cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is a penetrating peptide (IBW-002) used
 CC in the composition for mucosal vaccination using a counter ion and a
 CC penetrating peptide.

Seq Sequence 30 AA;

Query Match 95.8%; Score 114; DB 9; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.2e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDYITLALAGICQSAARLVQOLA 24
 Db 1 NYDYITLALAGICQSAARLVQOLA 23

RESULT 14

AEB08263 standard; protein; 213 AA.

XX

AC AEB08263;
 XX
 XX 25-AUG-2005 (first entry)
 DE Escherichia coli protein, SEQ ID NO: 61.
 XX
 XX Pharmacological, therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX antidiabetic disorder; endocrine disease; gastrointestinal disease;
 XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
 XX neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 XX dementia; multiple sclerosis; immune disorder; Huntington chorea;
 XX anticonvulsant; genetic disorder; cardiovascular disease;
 XX cardiovascular-gen.; atherosclerosis; arteriosclerosis;
 XX coronary artery disease; cardiac; vasotrophic; obesity; anorectic;
 XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 XX genitourinary disease; hematological disease; antianemic; anemia;
 XX autoimmune disease; immunosuppressive; immune deficiency;
 XX immunostimulant; infectious disease; antimicrobial; infection;
 XX erectile dysfunction; andrology; major depressive disorder;
 XX antidepressant; psychiatric disorder; pain; analgesic;
 XX bacterial infection; antibacterial; viral infection; virostatic;
 XX fungal infection; fungicide; parasitic infection; antiparasitic;
 XX renal failure; antifertility; antipneumatic; cytostatic;
 XX antineoplastic; hepatotropic; hepatitis B virus infection.
 XX Escherichia coli.
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHEN/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX
 XX MPI; 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Disclosure; SEQ ID NO 61; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for trans epithelial delivery of effector. The
 XX invention comprises the effector sequentially coupled with a counter ion
 XX and at least one hydrophobic agent, where the effector is selectively
 XX encapsulated into a complex. The invention is useful for translocating
 XX effectors across a biological barrier such as epithelial cells and
 XX endothelial cells during treatment and prevention of disease or
 XX pathological conditions (including endocrine disorders, diabetes;
 XX infertility, hormone deficiencies, osteoporosis, ophthalmological,
 XX disorders, neurodegenerative disorders, Alzheimer's disease,
 XX Parkinson's disease, multiple sclerosis, Huntington's disease,
 XX cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 XX coagulable states, coronary disease, cerebrovascular events, metabolic
 XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 XX hematological disorders, anemia of different entities, immunologic and
 XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
 XX infectious diseases, viral infections, bacterial infections, fungal
 XX infections, parasitic infections, neoplastic diseases, multi-factorial
 XX disorders, impotence, chronic pain, depression, different fibrosis states
 XX and short stature) and for mucosal vaccination against anthrax and
 XX hepatitis B. The present sequence is the Escherichia coli protein

CC containing penetrating peptide at N-terminal end.
 XX
 XX SQ Sequence 213 AA;

Query Match 95.8%; Score 114; DB 9; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQARLVQOLA 24
 |||||
 DB 4 NYDITLALAGICQARLVQOLA 26

RESULT 15

AB061272
 ID AB061272 standard; protein; 217 AA.

AC AB061272;

DT 29-JUL-2004 (first entry)

DE Klebsiella pneumoniae polypeptide seqid 7789.

KW Recombinant expression vector; transcription regulatory element;

KW Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX Klebsiella pneumoniae.

OS US6610836-B1.

PN 26-AUG-2003.

PP 27-JAN-2000; 2000US-00489039.

PR 29-JAN-1999; 99US-0117747P.

PA (GENO-) GENOME THERAPEUTICS CORP.

PI Breton GL, Osborne M;

DR MPI; 2003-895346/82.

XX N-PSDB; ACH94823.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 XX preparing a vaccine composition against Klebsiella pneumoniae.

PS Disclosure; SEQ ID NO 7789; 932pp; English.

CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention

XX Sequence 217 AA;

SQ

Query Match 92.4%; Score 110; DB 7; Length 217;

Best Local Similarity 91.3%; Pred. No. 9.8e-10;

Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQARLVQOLA 24
 |||||

DB 8 NYDITLALAGICQARLVQOLA 30

Search completed: January 23, 2006, 10:28:33

Job time : 79.6507 secs

This Page Marked for
Removal from File

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:52:43 ; Search time 55.1266 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-25

Perfect score: 119

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	24	4	US-10-665-184-25 Sequence 25, Appl
2	119	100.0	24	5	US-10-942-300-25 Sequence 25, Appl
3	119	100.0	24	5	US-10-501-838a-25 Sequence 25, Appl
4	119	100.0	31	4	US-10-665-184-35 Sequence 35, Appl
5	119	100.0	31	5	US-10-942-300-35 Sequence 35, Appl
6	119	100.0	31	5	US-10-501-838a-35 Sequence 35, Appl
7	114	95.8	23	4	US-10-665-184-3 Sequence 3, Appl1
8	114	95.8	23	5	US-10-942-300-3 Sequence 3, Appl1
9	114	95.8	23	5	US-10-501-838a-3 Sequence 3, Appl1
10	114	95.8	25	5	US-10-501-838a-19 Sequence 19, Appl
11	114	95.8	25	5	US-10-501-838a-23 Sequence 23, Appl
12	114	95.8	26	5	US-10-501-838a-21 Sequence 21, Appl
13	114	95.8	29	4	US-10-501-838a-18 Sequence 22, Appl
14	114	95.8	30	5	US-10-665-184-22 Sequence 22, Appl
15	114	95.8	30	5	US-10-942-300-22 Sequence 22, Appl
16	114	95.8	30	5	US-10-501-838a-20 Sequence 20, Appl
17	114	95.8	30	5	US-10-501-838a-22 Sequence 22, Appl
18	114	95.8	213	4	US-10-665-184-61 Sequence 61, Appl
19	114	95.8	213	5	US-10-942-300-61 Sequence 61, Appl
20	114	95.8	213	5	US-10-501-838a-55 Sequence 55, Appl
21	114	95.8	23	4	US-10-665-184-26 Sequence 26, Appl
22	114	95.8	23	5	US-10-942-300-26 Sequence 26, Appl
23	114	95.8	23	5	US-10-501-838a-26 Sequence 26, Appl
24	114	95.8	30	5	US-10-665-184-36 Sequence 36, Appl
25	114	95.8	30	5	US-10-942-300-36 Sequence 36, Appl
26	114	95.8	30	5	US-10-501-838a-36 Sequence 36, Appl
27	114	95.8	23	4	US-10-665-184-2 Sequence 2, Appl1

28	103	86.6	23	5	US-10-942-300-2 Sequence 2, Appl1
29	103	86.6	23	5	US-10-501-838a-2 Sequence 2, Appl1
30	103	86.6	203	4	US-10-665-184-60 Sequence 60, Appl
31	103	86.6	203	5	US-10-942-300-60 Sequence 60, Appl
32	103	86.6	203	5	US-10-501-838a-54 Sequence 54, Appl
33	101	84.9	22	4	US-10-665-184-27 Sequence 27, Appl
34	101	84.9	22	5	US-10-942-300-27 Sequence 27, Appl
35	101	84.9	22	5	US-10-501-838a-27 Sequence 27, Appl
36	101	84.9	29	4	US-10-665-184-37 Sequence 37, Appl
37	101	84.9	29	5	US-10-942-300-37 Sequence 37, Appl
38	101	84.9	29	5	US-10-501-838a-37 Sequence 37, Appl
39	98	82.4	23	4	US-10-665-184-29 Sequence 29, Appl
40	98	82.4	23	5	US-10-942-300-29 Sequence 29, Appl
41	98	82.4	23	5	US-10-501-838a-29 Sequence 29, Appl
42	98	82.4	29	4	US-10-665-184-33 Sequence 33, Appl
43	98	82.4	29	5	US-10-942-300-33 Sequence 33, Appl
44	98	82.4	30	5	US-10-501-838a-33 Sequence 33, Appl
45	95	79.8	23	4	US-10-665-184-1 Sequence 1, Appl1

ALIGNMENTS

```

RESULT 1
US-10-665-184-25
; Sequence 25, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-665-184-25
Query Match          100.0%; Score 119; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNYDITLALAGICOSARLVQOLA 24
DB 1 MNYDITLALAGICOSARLVQOLA 24
RESULT 2
US-10-942-300-25
; Sequence 25, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

```

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 25
;; LENGTH: 24
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-10-942-300-25

Query Match 100.0%; Score 119; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 6,3e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMYDITLALAGICQSARLVQQLA 24
Db 1 MMYDITLALAGICQSARLVQQLA 24

RESULT 3
US-10-501-838a-25
; Sequence 25, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-501-838a-25

Query Match 100.0%; Score 119; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 6,3e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMYDITLALAGICQSARLVQQLA 24
Db 1 MMYDITLALAGICQSARLVQQLA 24

RESULT 4
US-10-665-184-35
; Sequence 35, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial

;; FEATURE:
;; OTHER INFORMATION: Penetrating peptide
;; FEATURE:
;; NAME/KEY: MOD_RES
;; LOCATION: (1)-(1)
;; OTHER INFORMATION: ACETYLATION
;; FEATURE:
;; NAME/KEY: PEPTIDE
;; LOCATION: (31)-(31)
;; OTHER INFORMATION: wherein Xaa is Lysine-NH2
US-10-665-184-35

Query Match 100.0%; Score 119; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 8,3e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMYDITLALAGICQSARLVQQLA 24
Db 1 MMYDITLALAGICQSARLVQQLA 24

RESULT 5
US-10-942-300-35
; Sequence 35, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Acylated Penetrating Peptide
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (31)-(31)
; OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated
US-10-942-300-35

Query Match 100.0%; Score 119; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 8,3e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMYDITLALAGICQSARLVQQLA 24
Db 1 MMYDITLALAGICQSARLVQQLA 24

RESULT 6
US-10-501-838a-35
; Sequence 35, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL

```
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 35
/ LENGTH: 31
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic: penetrating peptide
/ NAME/KEY: MISC FEATURE
/ LOCATION: (28)..(28)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (31)..(31)
/ OTHER INFORMATION: groups of the lysine residue
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (31)..(31)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (31)..(31)
/ OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
/ OTHER INFORMATION: peptide via the free amino groups of the lysine residue
US-10-501-838A-35
```

```
Query Match          100.0%; Score 119; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 8,3e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 NYNYDITLALAGICQSARLVQOLA 24
Db 1 NYNYDITLALAGICQSARLVQOLA 24
```

```
RESULT 7
US-10-665-184-3
/ Sequence 3, Application US/10665184
/ Publication No. US20040146549A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-501CIP
/ CURRENT APPLICATION NUMBER: US/10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
US-10-665-184-3
```

```
Query Match          95.8%; Score 114; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 3,9e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 NYNYDITLALAGICQSARLVQOLA 24
Db 1 NYNYDITLALAGICQSARLVQOLA 23
```

```
RESULT 8
US-10-942-300-3
/ Sequence 3, Application US/10942300
/ Publication No. US20050136103A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-503
/ CURRENT APPLICATION NUMBER: US/10/942,300
/ PRIOR FILING DATE: 2004-09-16
/ PRIOR APPLICATION NUMBER: 10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 10/664,989
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 60/503,615
/ PRIOR FILING DATE: 2003-09-17
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
US-10-942-300-3
```

```
Query Match          95.8%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3,9e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 NYNYDITLALAGICQSARLVQOLA 24
Db 1 NYNYDITLALAGICQSARLVQOLA 23
```

```
RESULT 9
US-10-501-838A-3
/ Sequence 3, Application US/10501838A
/ Publication No. US20050215478A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel A.
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-501 NATL
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ PRIOR FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
US-10-501-838A-3
```

```
Query Match          95.8%; Score 114; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3,9e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 NYNYDITLALAGICQSARLVQOLA 24
Db 1 NYNYDITLALAGICQSARLVQOLA 23
```

```
RESULT 10
US-10-501-838A-19
/ Sequence 19, Application US/10501838A
/ Publication No. US20050215478A1
/ GENERAL INFORMATION:
```

APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT APPLICATION NUMBER: US/10/501,838A
PRIOR FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 19
LENGTH: 25
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic: penetrating peptide
NAME/KEY: MISC FEATURE
LOCATION: (25)-(25)
OTHER INFORMATION: wherein recombinant human insulin is coupled to the penetrating
US-10-501-838A-19

Query Match 95.8%; Score 114; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDITLALAGICQSARLVQOLA 24
Db 1 NYDITLALAGICQSARLVQOLA 23

RESULT 11
US-10-501-838A-23
Sequence 23, Application US/10501838A
Publication No. US20050215478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT APPLICATION NUMBER: US/10/501,838A
PRIOR FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 23
LENGTH: 25
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic: penetrating peptide
NAME/KEY: MISC FEATURE
LOCATION: (25)-(25)
OTHER INFORMATION: wherein the penetrating peptide is coupled to a linearized
OTHER INFORMATION: insulin receptor, which is in turn coupled to recombinant human
US-10-501-838A-23

Query Match 95.8%; Score 114; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDITLALAGICQSARLVQOLA 24
Db 1 NYDITLALAGICQSARLVQOLA 23

Db 1 NYDITLALAGICQSARLVQOLA 23

RESULT 12
US-10-501-838A-21
Sequence 21, Application US/10501838A
Publication No. US20050215478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT APPLICATION NUMBER: US/10/501,838A
PRIOR FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 21
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic: penetrating peptide
NAME/KEY: MISC FEATURE
LOCATION: (26)-(26)
OTHER INFORMATION: wherein heparin is coupled to the penetrating peptide via the
US-10-501-838A-21

Query Match 95.8%; Score 114; DB 5; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NYDITLALAGICQSARLVQOLA 24
Db 1 NYDITLALAGICQSARLVQOLA 23

RESULT 13
US-10-501-838A-18
Sequence 18, Application US/10501838A
Publication No. US20050215478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501 NATL
CURRENT APPLICATION NUMBER: US/10/501,838A
PRIOR FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 18
LENGTH: 29
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic: penetrating peptide
NAME/KEY: MISC FEATURE
LOCATION: (26)-(29)
OTHER INFORMATION: cleavable linker peptide
NAME/KEY: MISC FEATURE

LOCATION: (26)..(29)
OTHER INFORMATION: wherein recombinant human insulin is coupled to the penetrating
OTHER INFORMATION: peptide via the cleavable linker peptide
US-10-501-838a-18

Query Match 95.8%; Score 114; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 5e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQSARLVQOLA 24
DB 1 NYDITLALAGICQSARLVQOLA 23

RESULT 14
US-10-665-184-22

Sequence 22, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
APPLICANT: Cohen, Binat
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
TITLE OF INVENTION: Biological Barrier
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 22
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Penetrating peptide
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)..(1)
OTHER INFORMATION: ACETYLATION
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (30)..(30)
OTHER INFORMATION: wherein Xaa is Lysine-NH2
US-10-665-184-22

Query Match 95.8%; Score 114; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQSARLVQOLA 24
DB 1 NYDITLALAGICQSARLVQOLA 23

RESULT 15
US-10-942-300-22

Sequence 22, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
APPLICANT: Cohen, Binat
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
TITLE OF INVENTION: Biological Barrier
FILE REFERENCE: 24348-503
CURRENT APPLICATION NUMBER: US/10/942,300
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 60/503,615
PRIOR FILING DATE: 2003-09-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 22
LENGTH: 30

TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Acylated Penetrating Peptide
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (30)..(30)
OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated
US-10-942-300-22

Query Match 95.8%; Score 114; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NYDITLALAGICQSARLVQOLA 24
DB 1 NYDITLALAGICQSARLVQOLA 23

Search completed: January 23, 2006, 12:09:00
Job time : 55.1266 secs

THE NEW YORK PUBLIC LIBRARY
ASTOR LENOX TILDEN FOUNDATION
1911

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:32:56 ; Search time 12.5546 seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-26
Perfect score: 113
Sequence: 1 MYDITLALAGICOSARLVQOLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	108	95.6	208	2 AH0199 conserved hypochet
2	108	95.6	213	2 S19211 ycfC protein - Esc
3	108	95.6	213	2 D90829 hypothetical prote
4	108	95.6	213	2 B85687 hypothetical prote
5	105	92.9	215	2 AB0647 conserved hypochet
6	89	78.8	205	2 164155 hypothetical prote
7	80	70.8	211	2 D84960 hypothetical prote
8	72	63.7	205	2 B82237 conserved hypochet
9	46	40.7	335	2 S70671 lipopolysaccharide
10	45	39.8	188	2 J01065 coat protein - bee
11	45	39.8	947	2 AH0891 adeny1-transferase
12	44	38.9	204	2 E82680 conserved hypochet
13	44	38.9	211	2 T00567 hypothetical prote
14	43	38.1	280	2 A10326 cel operon repres
15	43	38.1	282	2 E71543 hypothetical prote
16	43	38.1	299	2 B81698 conserved hypochet
17	43	38.1	759	2 D70422 cellulose synthase
18	43	38.1	847	2 G95843 conserved hypochet
19	42.5	37.6	510	2 D96741 hypothetical prote
20	42	37.2	335	2 G71954 lipopolysaccharide
21	42	37.2	340	2 G64554 lipopolysaccharide
22	42	37.2	384	2 I57957 neurokinin 2 recep
23	42	37.2	384	2 S20303 neurokinin 2 recep
24	42	37.2	390	2 A46737 hypothetical prote
25	42	37.2	429	2 T24922 chitinase A (impor
26	41.5	36.7	1215	2 T22503 hypothetical prote
27	41	36.3	98	2 B91207 heat shock protein
28	41	36.3	144	2 B91207 heat shock protein
29	41	36.3	144	2 G65170 heat shock protein

30	41	36.3	144	2 D86053 heat shock protein
31	41	36.3	260	2 A34178 bacteriorhodopsin
32	41	36.3	274	2 S76154 hypothetical prote
33	41	36.3	317	2 AF0973 lipopolysaccharide
34	41	36.3	317	2 A44156 lipopolysaccharide
35	41	36.3	319	2 A40619 lipopolysaccharide
36	41	36.3	330	2 D86038 heptosyl transfera
37	41	36.3	330	2 C91191 neurokinin 2 recep
38	41	36.3	384	1 S00516 neurokinin 2 recep
39	41	36.3	398	1 J01059 neurokinin 2 recep
40	41	36.3	402	2 I56595 conserved hypochet
41	41	36.3	455	2 F82345 conserved hypochet
42	41	36.3	466	2 T30040 hypothetical prote
43	41	36.3	646	2 H88637 protein F53H.1 (I
44	41	36.3	1273	2 AF2421 hypothetical prote
45	40.5	35.8	248	2 AH2339 hypothetical prote

ALIGNMENTS

RESULT 1

AH0199 conserved hypothetical protein YP01637 [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #ext_change 09-Jul-2004

C/Accession: AH0199
R/Parikhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.

demo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:1158360

A/Accession: AH0199
A/Status: preliminary

A/Molecule type: DNA
A/Residues: 1-208 <NR>

A/Cross-references: UNIPROT:Q8ZF06; UNIPARC:UPI00000CD828; GB:AL590842; PIDN:CAC90459.1,
C/Genetics:

A/Gene: YP01637

Query Match 95.6% Score 108; DB 2; Length 208;
Best Local Similarity 100.0% Pred. No. 2.9e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YYDITLALAGICOSARLVQOLA 23
DB 5 YYDITLALAGICOSARLVQOLA 26

RESULT 2

S19211 ycfC protein - Escherichia coli (strain K-12)

C/Species: Escherichia coli
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #ext_change 09-Jul-2004

C/Accession: S19211; A64858
R/Green, S.M.; Drablie, W.T.

submitted to the EMBL Data Library, May 1991
A/Description: Molecular analysis of the purB-phop region of Escherichia coli K12.

A/Reference number: S19210
A/Accession: S19211

A/Status: preliminary
A/Molecule type: DNA

A/Residues: 1-213 <GRB>
A/Cross-references: UNIPROT:P25746; UNIPARC:UPI000013A673; EMBL:X59307; NID:G42582; PIDN:

A/Experimental source: UNIPROT:K-12
R/Blaetner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503

A/Accession: A64858

A/Status: nucleic acid sequence not shown; translation not shown

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <STO>
A:Cross-references: UNIPARC:UPI000005B4FF; GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: yfcJ; BU262

Query Match 70.8%; Score 80; DB 2; Length 211;
Best Local Similarity 89.5%; Pred. No. 1.3e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 YDITLALGICQSARLVQOLA 23
DB 7 YDITLALGICQSARLVQOLA 25

RESULT 8
B82237
conserved hypothetical protein VC1127 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: B82237 Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardron, D.; Ermolaeva, M.D.; Vamathevan, J.; Babs, S.; Qin, H.; Dragol, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: B82237
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-205 <HEI>
A:Cross-references: UNIPROT:Q9KX9; UNIPARC:UPI00000C2EB4; GB:AE004193; GB:AE003852; NID: A:Experimental source: serogroup O1, strain N16961, biotype El Tor
C:Genetics:
A:Gene: VC1127
A:Map position: 1

Query Match 63.7%; Score 72; DB 2; Length 205;
Best Local Similarity 66.7%; Pred. No. 0.00027;
Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDITLALGICQSARLVQOLA 23
DB 6 YDITLALGICQSARLVQOLA 26

RESULT 9
S70671
lipopolysaccharide heptosyltransferase (EC 2.4.99.-) I rFAC [similarity] - Bordetella pertussis
C:Species: Bordetella pertussis
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S70671 R.Allen, A.; Maskell, D.
Mol. Microbiol. 19, 37-52, 1996
A>Title: The identification, cloning and mutagenesis of a genetic locus required for lipopolysaccharide heptosyltransferase
A:Reference number: S70669; MUID:96419162; PMID:8821935
A:Accession: S70671
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-335 <ALL>
A:Cross-references: UNIPROT:Q45375; UNIPARC:UPI00000A5930; EMBL:X90711; NID:9992967; PID: A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
C:Genetics:
A:Gene: rFAC
C:Superfamily: heptosyltransferase I
C:Keywords: glycosyltransferase

Query Match 40.7%; Score 46; DB 2; Length 335;
Best Local Similarity 47.4%; Pred. No. 8.8;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 YDITLALGICQSARLVQOL 21
DB 89 YDITLALGICQSARLVQOL 107

RESULT 10
JCI065
coat protein - beet necrotic yellow vein mosaic virus
C:Species: Beet necrotic yellow vein mosaic virus
C>Date: 02-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C:Accession: JCI065 R.Yao, H.J.; Liu, Y.; Yu, J.L.; Cai, Z.N.; Ao, G.M.; Yang, L.L.
Chinese J. Biotechnol. 9, 147-151, 1993
A>Title: The cloning and sequencing of coat protein gene from beet necrotic yellow vein virus
A:Reference number: JCI065
A:Accession: JCI065
A:Molecule type: mRNA
A:Residues: 1-188 <YAO>
A:Cross-references: UNIPROT:Q86757; UNIPARC:UPI00000F0F13
C:Keywords: coat protein

Query Match 39.8%; Score 45; DB 2; Length 188;
Best Local Similarity 50.0%; Pred. No. 7.4;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 YDITLALGICQSARLV 18
DB 143 YDITLALGICQSARLV 158

RESULT 11
AH0891
adenyl-transferase [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C:Species: Salmonella enterica subsp. enterica serovar Typhimurium
A:Note: this species has also been called Salmonella typhimurium
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0891 R.Pearkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, F.; Davies, K.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
A:Reference number: AB0502; MUID:2154947; PMID:11677608
A:Accession: AH0891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-947 <PAR>
A:Cross-references: UNIPARC:UPI000005A4A4; GB:AL513382; PIDN:CAD07726.1; PID:G16504278; C:Genetics:
A:Gene: STY3380

Query Match 39.8%; Score 45; DB 2; Length 947;
Best Local Similarity 53.3%; Pred. No. 35;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 8 ALAGVCDASIMREL 22
DB 76 ALAGVCDASIMREL 90

RESULT 12
E82680
conserved hypothetical protein XF1439 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82680 R.anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: AB2515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number. A59328 below
A:Accession: E82680

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <SIM>
A:Cross-references: UNIPROT:O9PDE0, UNIPARC:UPI00000CC2723, GB:AE003974, GB:AE003849, NID
A:Experimental source: strain 945C
R:Simpon, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briotes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carver, H
de-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurama, E.E.; Laig
Chado, M.A.; Madeira, A.M.B.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsumura, A.Y.; Menck, C.F.M.; Miraca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsubako, M.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XFL439

Query Match 38.9%; Score 44; DB 2; Length 204;
Best Local Similarity 52.9%; Pred. No. 12;
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 LALAGCOSARLVQOLA 23
DB 10 LALAGVQALQVQROIA 26

RESULT 13

T00567
hypothetical protein Atc2g39570 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F12L6.23
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C:Accession: T00567; A84819
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC F12L6 genomic sequence.

A:Reference number: Z14168

A:Status: translated from GB/EMBL/DBJ

A:Accession: T00567

A:Molecule type: DNA
A:Residues: 1-411 <ROU>

A:Cross-references: UNIPROT:O80644; UNIPARC:UPI00000A2464; EMBL:AC004218; NID:93355463;

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

gus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-766, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: A84819

A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-411 <STO>

A:Cross-references: UNIPARC:UPI00000A2464; GB:AE002093; NID:93355486; PIDN:AA027848.1; C
C:Genetics:
A:Gene: Atc2g39570; F12L6.23

QY 1 MYDITLAL--AGIC 13

DB 346 VFYDVTALAKSIGIC 360

Query Match 38.9%; Score 44; DB 2; Length 411;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 9; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

RESULT 14

AI0326
cel operon repressor [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AI0326

R:Parkhill, J.; Wen, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarrag, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AI0326

A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-280 <KUR>

A:Cross-references: UNIPROT:Q82DA7; UNIPARC:UPI00000DCAB8; GB:AL590842; PIDN:CAC92920.1;
C:Genetics:
A:Gene: celd

Query Match 38.1%; Score 43; DB 2; Length 280;
Best Local Similarity 58.3%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YYDITLALAGIC 13
DB 40 YYEYTLITLITGMC 51

RESULT 15

E71543
hypothetical protein CT203 - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999

C:Accession: E71543

R:Stephens, R.S.; Kaiman, S.; Lammell, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac

A:Reference number: A71570; MUID:99000809; PMID:9784136

A:Accession: E71543

A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-282 <ARN>

A:Cross-references: UNIPARC:UPI0000179B22; GB:AE001294; GB:AE001273; NID:93326609; PIDN:
A:Experimental source: serotype D, strain UW-3/Cx

QY 4 DITLALAGCOSARLVQOL 22
DB 133 DLSRIILASVCDDVALIOEL 151

Search completed: January 23, 2006, 10:43:09
Job time : 12.5546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:31:19 ; Search time 73.3188 Seconds
(without alignments)
221.323 Million cell updates/sec

Title: US-10-501-838a-26
Perfect score: 113
Sequence: 1 MYDITLALAGICGASRLVQOLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_05.80: *
2: uniprot_sprot: *
3: uniprot_trembl: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	108	95.6	208 1 Y1637_YERPE	082f66 Yersinia pe
2	108	95.6	208 1 Q669Q3_YERPS	0669q3 Yersinia ps
3	108	95.6	213 1 YCPC_ECO57	08x736 escherichia
4	108	95.6	213 1 YCPC_ECOL6	08f1b7 escherichia
5	108	95.6	213 1 YCPC_ECOL1	P25766 escherichia
6	108	95.6	213 1 YCPC_SHIFL	0831f8 shigella fl
7	105	92.9	213 1 YCPC_SALTY	08z7n0 salmonella
8	105	92.9	213 1 YCPC_SALTY	08z7n0 salmonella
9	105	92.9	213 2 Q5PMJ3_SALPA	Q5pmj3 salmonella
10	101	89.4	215 2 Q57QCL_SALGH	Q57qcl salmonella
11	100	88.5	204 2 Q65VVS_MANSN	Q65vvs manneinia
12	99	87.6	204 2 Q6DAE8_ERWCT	Q6dae8 erwania car
13	97	85.8	203 1 Y1850_PASWU	Q9c3f8 pasteurella
14	91	80.5	208 1 Y2805_PROHL	Q7n3b4 photorhabdu
15	89	78.8	205 1 Y638_HABIN	P47366 haemophilus
16	89	78.8	205 2 Q4QMS9_HABE8	Q4qms9 haemophilus
17	81	71.7	215 1 Y1650_HAEPU	Q7v135 haemophilus
18	80	70.8	211 1 Y262_EUCAT	P57350 buchnera ap
19	72	63.7	205 1 Y1127_VIBCH	Q9kx99 vibrio chol
20	72	63.7	205 1 Y1129_VIBPA	Q87qno vibrio para
21	70	61.9	205 1 Y1342_VIBVU	Q87ml5 vibrio vuln
22	70	61.9	205 1 Y2927_VIBVU	Q8d8r4 vibrio vuln
23	69	61.1	205 2 Q5E3W6_VIBF1	Q5e3w6 vibrio fisc
24	68	60.2	217 1 Y243_BUCBP	Q9s9e4 buchnera ap
25	62	54.9	205 2 Q6LTI9_PHOPR	Q6lti9 photobacter
26	50	44.2	3223 2 Q5K5U8_ECOL1	Q5k5u8 escherichia
27	50	44.2	3223 2 Q5K5U8_ECOL1	Q5k5u8 escherichia
28	50	44.2	3223 2 Q8VNR6_ECOL1	Q8vnr6 escherichia
29	50	44.2	3223 2 Q8VNR6_ECOL1	Q8vnr6 escherichia
30	50	44.2	3223 2 Q9RM48_ECOL1	Q9rm48 escherichia
31	50	44.2	3223 2 Q9RPH1_ECOL1	Q9rph1 escherichia

32	50	44.2	3223 2 Q5MBA6_ECOL1	Q5mba6 escherichia
33	49	43.4	216 2 Q6C8B5_YARLI	Q6c8b5 yarrowia li
34	49	43.4	266 2 Q7S4H4_NEUCR	Q7s4h4 neurospora
35	49	43.4	391 2 Q5GVQ3_XANOR	Q5gvq3 xanthomonas
36	49	43.4	3208 2 Q5DKN7_GENTR	Q5dkn7 enterobacter
37	48	42.5	197 2 Q9BER3_RHILQ	Q9ber3 rhizobium l
38	48	42.5	209 2 Q6CAB8_METCA	Q6cab8 methylococc
39	48	42.5	597 2 Q726U4_DRSVH	Q726u4 deulifeyldr
40	47	41.6	292 2 Q6LRP9_CABRR	Q6lrp9 caenorhabdi
41	46.5	41.2	489 2 Q5FSJ7_GLUOX	Q5fsj7 gluconobact
42	46	40.7	320 2 Q7W235_BORPA	Q7w235 bordetella
43	46	40.7	320 2 Q7WR13_BORBR	Q7wr13 bordetella
44	46	40.7	335 2 Q8B002_BORBR	Q8b002 bordetella
45	46	40.7	335 2 Q45375_BORPE	Q45375 bordetella

ALIGNMENTS

RESULT 1
Y1637_YERPE STANDARD; PRT; 208 AA.
AC Q8ZFQ6; Q8DQO1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein YP01637/Y1798/YP1767.
GN OrderedCusNames=YPO1637, Y1798, YP1767;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_Taxid=632;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parthill J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
RA Prentice M.B., Sebatina M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Rarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Skimonde M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,
RA "Genome sequence of Yersinia pestis, the causative agent of plague.",
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA DOI=10.1126/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G., Ilt, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
RA Perry R.D.,
RT "Genome sequence of Yersinia pestis KIM.",
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Chi B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.,
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
avirulent to humans".
RL DNA Res. 11:1179-1197(2004).
CC -I- SIMILARITY: Belongs to the UPF0274 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

```
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ414149; CAC90459.1; -; Genomic_DNA.
CC EMBL; AE013782; AAM8366.1; ALT_INIT; Genomic_DNA.
CC EMBL; AE017133; AAS61994.1; -; Genomic_DNA.
CC PIR; AH0199; AH0199.
CC SMR; Q8ZFO6; 2-207.
CC HAMAP; MF_00695; -; 1.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC CONFLICT 1 M -> MIV (in Ref. 2).
CC SEQUENCE 208 AA; 22730 MW; DQCC2B3BDE499723 CRC64;

Query Match 95.6%; Score 108; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 YYDTTLAAGICQSARLVQOLA 23
Db 5 YYDTTLAAGICQSARLVQOLA 26

RESULT 2
066903 YERPS PRELIMINARY; PRT; 208 AA.
AC 066903;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=YPRB2431;
OS Yersinia pseudotuberculosis.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Yersinia.
CC NCBI_TaxID=633;
CC [1]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=IP32953 / Serotype 1;
CC PubMed=15358858; DOI=10.1073/pnas.0404012101;
CC Chain P.S.G., Camfield E., Larimer F.W., Lamerdin J., Stoutland P.O.,
CC Regala W.M., Georgescu A.M., Verges L.M., Land M.L., Motin V.L.,
CC Brubaker R.R., Fowler J., Himebusch J., Marceau M., Medicine C.,
CC Simone M., Cheneal-Francois V., Souza B., Dacheux D., Elliott J.M.,
CC Derbise A., Hauser L.J., Garcia E.;
CC "Insights into the evolution of Yersinia pestis through whole-genome
CC comparison with Yersinia pseudotuberculosis."
CC Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
CC EMBL; BX936398; CAH21669.1; -; Genomic_DNA.
CC SMR; Q66903; 2-207.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC SEQUENCE 208 AA; 22730 MW; DQCC2B3BDE499723 CRC64;

Query Match 95.6%; Score 108; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 YYDTTLAAGICQSARLVQOLA 23
Db 5 YYDTTLAAGICQSARLVQOLA 26

RESULT 3
YCFC_ECOL57 STANDARD; PRT; 213 AA.
ID YCFC_ECOL57;
AC 08X736;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.

DE Hypothetical UPF0274 protein ycfC.
```

```
GN Name=ycfC; OrderedLocustNames=z1861, Ecet1604;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
CC NCBI_TaxID=83334;
CC [1]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
CC MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
CC Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
CC Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
CC Poefel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
CC Grobeck E.J., Davis N.W., Lim A., Dimmlana B.T., Potamousts K.,
CC Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
CC Welch R.A., Blattner F.R.;
CC "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
CC Nature 409:529-533(2001).
CC [2]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
CC MEDLINE=21156231; PubMed=11258796;
CC Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
CC Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
CC Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
CC Kubara S., Shiba T., Hattori M., Shinagawa H.;
CC "Complete genome sequence of enterohaemorrhagic Escherichia coli
CC O157:H7 and genomic comparison with a laboratory strain K-12."
CC J. Bact. 191:1618-1627(2001).
CC RL DNA Res. 8:11-22(2001).
CC CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC CC
CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB005174; AAC55958.1; -; Genomic_DNA.
CC EMBL; BA000007; BAB35027.1; -; Genomic_DNA.
CC PIR; B85687; B85687.
CC DR PIR; D90829; D90829.
CC SMR; Q8X736; 2-213.
CC DR HAMAP; MF_00695; -; 1.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC SEQUENCE 213 AA; 22947 MW; E26E9678C344E2 CRC64;

Query Match 95.6%; Score 108; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 2 YYDTTLAAGICQSARLVQOLA 23
Db 5 YYDTTLAAGICQSARLVQOLA 26

RESULT 4
YCFC_ECOL6 STANDARD; PRT; 213 AA.
ID YCFC_ECOL6;
AC Q8F1B7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocustNames=c1511;
OS Escherichia coli O6.
CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC Enterobacteriaceae; Escherichia.
CC NCBI_TaxID=217992;
CC [1]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
CC MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
```


RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raebio D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic *Escherichia coli*.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, AB016759; AAN79980.1; ALT_INT; Genomic_DNA.
CC SMR, Q8F1B7; 2-213.
DR HAMAP, MF_00695; -; 1.
DR InterPro, IPR007451; DUF489.
DR Pfam, PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22934 MW; 71C8F96F8BA37B CRC64;

Query Match 95.6%; Score 108; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YYDITLALAGICOSARLVQOLA 23
Db 5 YYDITLALAGICOSARLVQOLA 26

RESULT 5
YCFC_ECOLI STANDARD; PRT; 213 AA.
AC YCFC_ECOLI 225746;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC (ORF-23).
GN Name:ycfC; Ordered locus names=b1132;
OS *Escherichia coli*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Escherichia*.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12;
RA Green S.M., Drabble W.T.;
RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12;
RX MEDLINE=92104952; PubMed=1729205;
RT "Escherichia coli purr gene: cloning, nucleotide sequence, and
RT regulation by purr.";
RT J. Bacteriol. 174:130-136(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12 / MG1655;
RX MEDLINE=97446617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Klinkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RT Science 277:1453-1474(1997).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kasahimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saio N.,
RA Samesi G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiiuchi T.;
RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RT DNA Res. 3:137-155(1996).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, X59307; CAA1995.1; -; Genomic_DNA.
CC EMBL, M74924; -; NOT_ANNOTATED_CDS; Genomic_DNA.
CC EMBL, U00096; AAC74216.1; -; Genomic_DNA.
CC EMBL, D90748; BAA35954.1; -; Genomic_DNA.
CC EMBL, D90749; BAA35963.1; -; Genomic_DNA.
CC PIR, S19211; S19211.
DR PDB; 1Q24; X-ray; A=2-213.
DR PDB; 1SD1; X-ray; A=2-213.
DR EcoBASS; EB1321; -;
DR EcoGene; EG11345; ycfC.
DR HAMAP, MF_00695; -; 1.
DR InterPro, IPR007451; DUF489.
DR Pfam, PF04356; DUF489; 1.
KW 3D-structure; Complete proteome; Hypothetical protein.
SQ SEQUENCE 213 AA; 22948 MW; E265F9698C3CE542 CRC64;

Query Match 95.6%; Score 108; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YYDITLALAGICOSARLVQOLA 23
Db 5 YYDITLALAGICOSARLVQOLA 26

RESULT 6
YCFC_SHIFL STANDARD; PRT; 213 AA.
AC 083LF8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name:ycfC; Ordered locus names=SF1151, S1234;
OS *Shigella flexneri*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Shigella*.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
RT through comparison with genomes of *Escherichia coli* K12 and O157.";
RT Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
DOI=10.1128/JAI.71.5.2775-2786.2003;

```

RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Pountner G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RU Infect. Immun. 71:2775-2786(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB005674; AAN42768.1; ALT_INIT; Genomic_DNA.
CC EMBL; AB016982; AAP1657.1; -; Genomic_DNA.
CC SMR; Q83LFG; 2-213.
CC HAMAP; MF_00695; -; 1.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC KW SEQUENCE 213 AA; 22894 MW; 0B0BEC698C29FB4D CRC64;
SQ

Query Match          95.6%; Score 108; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YYDITLALAGICQSARLVQOLA 23
DB 5 YYDITLALAGICQSARLVQOLA 26

RESULT 7
YCFC_SALTI STANDARD; PRT; 213 AA.
AC 0827H0;
DT 10-OCT-2003 (Rel. 42, Last Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocNames=STY1273, t1687;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
CC NCB1_TaxID=601;
OK

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parthill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Fretwell T., Hamlin R., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Kelowe A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RU Nature 413:848-852(2001).
CC [2]
CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
CC STRAIN=Ty2 / ATCC 700931;
CC MEDLINE=22531367; PubMed=12644504;
CC DOI=10.1128/JB.185.7.2330-2337.2003;
CC Deng W., Itoh S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
CC Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.,
CC "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
CC and CT18.";
CC J. Bacteriol. 185:2330-2337(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----

```

```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AL627269; CAD08357.1; ALT_INIT; Genomic_DNA.
CC EMBL; AL016839; AAO69312.1; ALT_INIT; Genomic_DNA.
CC SMR; Q827H0; 2-213.
CC HAMAP; MF_00695; -; 1.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC KW SEQUENCE 213 AA; 22942 MW; 63192D1C8066ED58 CRC64;
SQ

Query Match          92.9%; Score 105; DB 1; Length 213;
Best Local Similarity 95.5%; Pred. No. 1e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YYDITLALAGICQSARLVQOLA 23
DB 5 YYDITLALAGICQSARLVQOLA 26

RESULT 8
YCFC_SALTY STANDARD; PRT; 213 AA.
AC 082P25;
DT 10-OCT-2003 (Rel. 42, Last Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocNames=STM1233;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
CC NCB1_TaxID=602;
OK

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA STRAIN=L72 / SGSC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanders K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT L72.";
RU Nature 413:852-856(2001).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB008754; AAL20162.1; ALT_INIT; Genomic_DNA.
CC SMR; Q82P25; 2-213.
CC StyGene; SGT7277; ycfC.
CC HAMAP; MF_00695; -; 1.
CC InterPro; IPR007451; DUF489.
CC Pfam; PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC KW SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;
SQ

Query Match          92.9%; Score 105; DB 1; Length 213;
Best Local Similarity 95.5%; Pred. No. 1e-08;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YYDITLALAGICQSARLVQOLA 23
DB 5 YYDITLALAGICQSARLVQOLA 26

```

RESULT 9

OSPMJ3_SALPA PRELIMINARY; PRT; 213 AA.

AC 05PMJ3; 01-FEB-2005 (TRENBLrel. 29, Created)

DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)

DE Hypothetical protein ycfC.

GN Name=yfcC; OrderedLocNames=SP1617;

OS Salmonella paratyphi-a.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=54388;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 9150;

RX PubMed=1531882; DOI=10.1038/ng1470;

RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,

RA Porrolik S., Sabo A., Meyer R., Bieri T., Ozeraky P., McElliott M.,

RA Hartline C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,

RA Kohlberg S., Strong C., Du P., Carter J., Kremicki C., Layman D.,

RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Mink P., Florea L.,

RA Delaunay K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,

RA Spiegh J., Wilson R.K.;

RT "Comparison of genome degradation in Paratyphi A and Typhi, human-

restricted serovars of Salmonella enterica that cause typhoid.";

RL Nat. Genet. 36:1268-1274(2004).

DR EMBL; CP000026; AAV7544.1; -; Genomic_DNA.

DR SMR; 05PMJ3; 2-213.

DR InterPro; IPR007451; DUF489.

DR Pfam; PF04356; DUF489; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match 92.9%; Score 105; DB 2; Length 213;

Best Local Similarity 95.5%; Pred. No. 1e-08;

Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YYDITLALAGICQSARLVQOLA 23

DB 5 YYDITLALAGICQSARLVQOLA 26

RESULT 10

OS7OC1_SALCH PRELIMINARY; PRT; 215 AA.

AC 057OC1; 10-MAY-2005 (TRENBLrel. 30, Created)

DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)

DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)

DE Hypothetical protein ycfC.

GN Name=yfcC; OrderedLocNames=SC1184;

OS Salmonella cholerae-suis (Salmonella enterica).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=591;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=SC-B67;

RX PubMed=15781495;

RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,

RA Wang H.-S., Lee Y.-S.;

RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a

highly invasive and resistant zoonotic pathogen.";

RL Nucleic Acids Res. 33:1690-1698(2005).

DR EMBL; AB017220; AAK5090.1; -; Genomic_DNA.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 215 AA; 23157 MW; C153820629F8B8D CRC64;

Query Match 89.4%; Score 101; DB 2; Length 215;

Best Local Similarity 95.2%; Pred. No. 4.6e-08;

Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YYDITLALAGICQSARLVQOL 22

DB 7 YYDITLALAGICQSARLVQOL 27

RESULT 11

OS6VVS_MANSM PRELIMINARY; PRT; 204 AA.

AC 065VVS; 25-OCT-2004 (TRENBLrel. 28, Created)

DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)

DE Hypothetical protein.

GN OrderedLocNames=MS0298;

OS Mannheimia succiniciproducens (strain MBL55B).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Mannheimia.

OX NCBI_TaxID=221988;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX PubMed=15378067; DOI=10.1038/nbt1010;

RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,

RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;

RT "The genome sequence of the capnophilic rumen bacterium Mannheimia

succiniciproducens.";

RL Nat. Biotechnol. 22:1275-1281(2004).

DR EMBL; AB016827; AAU36905.1; -; Genomic_DNA.

DR InterPro; IPR007451; DUF489.

DR Pfam; PF04356; DUF489; 1.

KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 204 AA; 23036 MW; 3D64A70A30C33CB CRC64;

Query Match 88.5%; Score 100; DB 2; Length 204;

Best Local Similarity 86.4%; Pred. No. 6.4e-08;

Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YYDITLALAGICQSARLVQOLA 23

DB 5 YYDITLALAGICQSARLVQOLA 26

RESULT 12

OS6D4B8_ERWCT PRELIMINARY; PRT; 212 AA.

AC 06D4B8; 25-OCT-2004 (TRENBLrel. 28, Created)

DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)

DE Hypothetical protein.

GN OrderedLocNames=SCA2443;

OS Erwilia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Pectobacterium.

OX NCBI_TaxID=29471;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=SCRI 1043 / ATCC BAA-672;

RX PubMed=15263089; DOI=10.1073/pnas.0402424101;

RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,

RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,

RA Akin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,

RA Fraser A., Hance Z., Haubert H., Jagels K., Moule S., Norbertczak H.,

RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,

RA Salmon D.G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;

RT "Genome sequence of the enterobacterial phytopathogen Erwilia

carotovora subsp. atroseptica and characterization of virulence

factors.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).

DR EMBL; BX950851; CAG75345.1; -; Genomic_DNA.

DR SMR; 06D4B8; 2-207

DR InterPro; IPR007451; DUF489.

DR Pfam: PF04356; DUF489; 1.
KW Complete proteome: Hypothetical protein.
SQ SEQUENCE 212 AA; 22892 MW; A67813005F328150 CRC64;

Query Match
Best Local Similarity 90.9%; Pred. No. 9, 7e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YYDTTLAAGICQSGRLVQQLA 23
Db 5 YVEITLALAGICQSGRLVQQLA 26

RESULT 13
Y1850_PASMU STANDARD; PRT; 203 AA.

AC Q9CJY8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein PM1850.
GN OrderedLocustNames=PM1850;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittem T.S., Kapur V.,
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AE006223; AK03934.1; -; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro: IPR007451; DUF489.
DR Pfam: PF04356; DUF489; 1.
KW Complete proteome: Hypothetical protein.
SQ SEQUENCE 203 AA; 22717 MW; 5EB91BFB0E181EDF CRC64;

Query Match
Best Local Similarity 85.8%; Score 97; DB 1; Length 203;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YYDTTLAAGICQSGRLVQQLA 23
Db 4 YYDTTLAAGICQSGRLVQQLA 25

RESULT 14
Y2805_PHOLL STANDARD; PRT; 208 AA.

AC Q7N3B4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein plu2805.
GN OrderedLocustNames=plu2805;
OS Photorhabdus luminescens (subsp. launondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IT01;

RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchand E., Rusluk C., Frangoul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Bouraux-Eude C., Chandler M., Charles J.-F.,
RA Daas E., Deroze S., Derzelle S., Freysinet G., Gaudriault S.,
RA Medigue C., Lancel A., Powell K., Siguler P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunet F.,
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens";
RL Nat. Biotechnol. 21:1307-1313(2003).

CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: BX571868; CAE15179.1; -; Genomic_DNA.
DR Photolysis; plu2805; -.
DR HAMAP; MF_00695; -; 1.
DR InterPro: IPR007451; DUF489.
DR Pfam: PF04356; DUF489; 1.
KW Complete proteome: Hypothetical protein.
SQ SEQUENCE 208 AA; 22831 MW; 3B484FA1FB836AD CRC64;

Query Match
Best Local Similarity 80.5%; Score 91; DB 1; Length 208;
Matches 18; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YYDTTLAAGICQSGRLVQQLA 23
Db 5 HYNITLALAGICQSGRLVQQLS 26

RESULT 15
Y638_HAEIN STANDARD; PRT; 205 AA.

AC P44796;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein HI0638.
GN OrderedLocustNames=HI0638;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=9530630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Pine L.D., Fritchman J.L., Fuhrmann J.L., Georgagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).

CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: U32747; AAC22298.1; -; Genomic_DNA.
DR PIR; I64155; I64155.

DR TIGR_H10638; -. 1.
DR HAMAP, MF_006951; -. 1.
DR InterPro, IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 2325 MW; 81F31B016BFB2DA CRC64;

Query Match	78.8%	Score 89	DB 1	Length 205
Best Local Similarity	77.3%	Pred. No. 4.1e-06		
Matches 17	Conservative 3	Mismatches 2	Indels 0	Gaps 0

Qy 2 YYDITLALAGICQSAKLVQQLA 23
|:| | | | | : | | | | |
Db 4 YHDI V L A L A G V C S A K L V H Q L A 25

Search completed: January 23, 2006, 10:40:57
Job time : 74.3188 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:35:50 (Search time 21.0917 Seconds
(without alignments)
90.156 Million cell updates/sec)

Title: US-10-501-838A-26

Sequence: 1 MYDITLALGICQSAALVQOLA 23

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/6.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/H.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/RS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/BACKfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	92.0	217	2	US-09-489-039A-7789
2	87	77.0	216	2	US-09-543-681A-7993
3	48	42.5	231	2	US-09-830-433A-32
4	45	39.8	138	2	US-09-270-767-58510
5	45	39.8	433	2	US-09-270-767-43172
6	43	38.1	231	2	US-09-830-433A-71
7	42.5	37.6	427	2	US-09-248-796A-17203
8	42.5	37.6	762	2	US-09-228-966-114
9	42.5	37.6	762	2	US-10-101-464A-114
10	42	37.2	369	1	US-07-937-609-19
11	42	37.2	369	2	US-08-029-170-19
12	42	37.2	369	2	US-09-443-745-19
13	42	37.2	823	2	US-09-253-991A-23655
14	41	36.3	60	2	US-09-248-796A-77407
15	41	36.3	114	2	US-09-902-540-10640
16	41	36.3	309	2	US-09-902-540-12828
17	41	36.3	384	1	US-08-103-170-10
18	41	36.3	387	1	US-08-196-989B-14
19	41	36.3	387	1	US-08-760-936-14
20	41	36.3	387	2	US-09-225-024-14
21	41	36.3	398	2	US-09-826-509-527
22	41	36.3	499	2	US-09-902-540-14780
23	41	36.3	588	2	US-09-605-703B-280
24	41	36.3	588	2	US-09-605-703B-282
25	41	36.3	741	2	US-09-252-991A-27062
26	40.5	35.8	717	2	US-09-107-532A-6350
27	40	35.4	169	2	US-09-489-039A-10685

28	40	35.4	311	2	US-09-489-039A-12640	Sequence 12640, A
29	40	35.4	324	2	US-09-489-039A-10801	Sequence 10801, A
30	40	35.4	398	2	US-09-543-681A-7046	Sequence 7046, Ap
31	40	35.4	409	1	US-08-360-673-6	Sequence 6, Appl
32	40	35.4	409	5	US-09-640-305-6	Sequence 6, Appl1
33	40	35.4	410	1	US-08-088-633-2	Sequence 2, Appl1
34	40	35.4	410	1	US-08-245-756-2	Sequence 2, Appl1
35	40	35.4	410	1	US-08-441-750-2	Sequence 2, Appl1
36	40	35.4	410	1	US-08-441-751-2	Sequence 2, Appl1
37	40	35.4	410	2	PCT-UB92-02521-2	Sequence 18383, A
38	40	35.4	440	2	US-09-248-796A-18383	Sequence 20352, A
39	40	35.4	576	2	US-09-248-796A-20352	Sequence 16773, A
40	40	35.4	1021	2	US-09-902-540-16773	Sequence 300, App
41	40	35.0	73	2	US-09-489-847-300	Sequence 229, App
42	39.5	35.0	74	2	US-09-489-847-229	Sequence 146, App
43	39.5	35.0	108	2	US-09-489-847-146	Sequence 3029, Ap
44	39.5	35.0	236	2	US-09-583-110-3029	Sequence 4815, Ap
45	39.5	35.0	241	2	US-09-107-433-4815	

ALIGNMENTS

```
RESULT 1
US-09-489-039A-7789
; Sequence 7789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7789
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match      92.0%; Score 104; DB 2; Length 217;
Best Local Similarity 90.9%; Pred. No. 1,4e+09;
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 YNDITLALGICQSAALVQOLA 23
Db      9 YNDITLALGVCQARLVQOLA 30

RESULT 2
US-09-543-681A-7993
; Sequence 7993, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7993
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7993

Query Match      77.0%; Score 87; DB 2; Length 216;
```

Best Local Similarity 77.3%; Pred. No. 1e-06;
Matches 17; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
US-09-830-433A-32
; Sequence 32, Application US/09830433A

```

GENERAL INFORMATION:
APPLICANT: AUTAMR et al.
TITLE OF INVENTION: Nucleic acids and polypeptides specific for pathogenic
TITLE OF INVENTION: strains of the Neisseria genus
FILE REFERENCE: P07180US00/BAS
CURRENT APPLICATION NUMBER: US/09/830,433A
CURRENT FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: FR 98 13 693
PRIOR FILING DATE: 1998-10-30
NUMBER OF SEQ ID NOS: 129
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 231
TYPE: PR1
ORGANISM: Neisseria meningitidis
US-09-830-433A-32

```

Query Match	42.5%	Score 48;	DB 2;	Length 231;
Best Local Similarity	52.9%;	Pred. No. 4.1;		
Matches	9;	Conservative	3;	Mismatches 5;
				Indels 0;
				Gaps 0;

```
QY      1 MYDYITLALAGICQ$AR 17
        :|||: |||: ||
DB      194 LYYDLASAPAPICRQ$AR 210
```

```

RESULT 4
US-09-270-767-58510
: Sequence 58510, Application US/09270767
: Patent No. 6703491
: GENERAL INFORMATION:
: APPLICANT: Homburger et al.
: TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
: FILE REFERENCE: File Reference: 732c-094
: CURRENT APPLICATION NUMBER: US/09/270,767
: CURRENT FILING DATE: 1999-03-17
: NUMBER OF SEQ ID NOS: 62517
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 58510
: LENGTH: 138
: TYPE: PRN
: ORGANISM: Drosophila melanogaster
US-09-270-767-58510

```

Query Match	39.8%;	Score 45;	DB 2;	Length 138;
Best Local Similarity	56.2%;	Pred. No. 7.2;		
Matches	9;	Conservative	2;	Mismatches 5;
				Indels 0;
				Gaps 0;

```
QY      2 YDITLALAGICQ SAR 17
          | | | : | | | : | |
Db      25 YADIEIALAGLADKAR 40
```

RESULT 5
 US-09-270-767-43172
 ; Sequence 43172, Application US/09270767
 ; Patent No. 7673491
 ; GENERAL INFORMATION:
 ; APPLICANT: Homburger et al.
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
 ; FILE REFERENCE: File Reference: 7326-094

```

: CURRENT APPLICATION NUMBER: US/09/270,767
:
: CURRENT FILING DATE: 1999-03-17
:
: NUMBER OF SEQ ID NOS: 62517
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 43172
:
: LENGTH: 433
:
: type: DPM

```

OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43172

Query Match	39.84;	Score 45;	DB 2;	Length 433;
Best Local Similarity	56.24;	Pred. No. 27;		
Matches	9;	Conservative	2;	Mismatches 5; Indels 0; Gaps 0;

```
QY      2 YYDITLALAGICQ$AR 17
          | | | : | | | : | |
DB      320 YADIBIALAGLADKAR 335
```

RESULT 6
US-09-830-433A-71
; Sequence 71, Application US/09830433A

```

? GENERAL INFORMATION:
? APPLICANT: AUTAME et al.
? TITLE OF INVENTION: Nucleic acids and polypeptides specific for pathogenic
? TITLE OF INVENTION: strains of the Neisseria genus
? FILE REFERENCE: P07180US00/BAS
? CURRENT APPLICATION NUMBER: US/09/830,433A
? CURRENT FILING DATE: 2001-04-26
? PRIOR APPLICATION NUMBER: FR 98 13 653
? PRIOR FILING DATE: 1998-10-30
? NUMBER OF SEQ ID NOS: 129
? SOFTWARE: PatentIn Ver. 2.1

```

```

; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-09-830-433A-71

```

Query Match	38.1%	Score 43;	DB 2;	Length 231;
Best Local Similarity	41.2%	Pred. No. 28;		
Matches	7;	Conservative	3;	Mismatches 7;
				Indels 0;
				Gaps 0;

Qy 1 MYDYITLALAGICQ\$AR 17
:|||: |||:
Db 194 LYDYDLAFRAGTICRQ\$AR 210

RESULT 7
US-09-248-796A-17203
; Sequence 17203, Application US/09248796A
; Patent No. 6747137

APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

CURRENT APPLICATION NUMBER: US 60/074,725
 CURRENT FILING DATE: 1998-02-12
 PRIOR APPLICATION NUMBER: US 60/096,409
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 28208
 SEQ ID NO 17203

```

; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
;

```


LOCATION: (339)
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-17203

Query Match 37.6%; Score 42.5; DB 2; Length 427;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 14; Conservative 1; Mismatches 6; Indels 7; Gaps 2;

QY 3 YDITLALAG--IC-----GSARLVQOLA 23
DB 120 YDLKTTTGVPCGCGADQASALVQOLA 147

RESULT 8
US-09-228-986-114
Sequence 114, Application US/09228986

PATENT No. 6359198
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Niels
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
FILE REFERENCE: 11000/1020
CURRENT APPLICATION NUMBER: US/09/228,986
CURRENT FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 130
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 114
LENGTH: 762
TYPE: PRT
ORGANISM: Pinus radiata
US-09-228-986-114

Query Match 37.6%; Score 42.5; DB 2; Length 762;
Best Local Similarity 70.6%; Pred. No. 1.3e+02;
Matches 12; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 8 ALAGICQASRL-VQOLA 23
DB 387 ALSIILQASRLNVEOLA 403

RESULT 9
US-10-101-464A-114
Sequence 114, Application US/10101464A

PATENT No. 6768041
GENERAL INFORMATION:
APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
FILE REFERENCE: 11000.1020C2
CURRENT APPLICATION NUMBER: US/10/101,464A
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724
PRIOR FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 114
LENGTH: 762
TYPE: PRT
ORGANISM: Pinus radiata
US-10-101-464A-114

Query Match 37.6%; Score 42.5; DB 2; Length 762;
Best Local Similarity 70.6%; Pred. No. 1.3e+02;

Matches 12; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 8 ALAGICQASRL-VQOLA 23
DB 387 ALSIILQASRLNVEOLA 403

RESULT 10
US-07-937-609-19
Sequence 19, Application US/07937609
PATENT No. 5319073
GENERAL INFORMATION:

APPLICANT: WANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: 1linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-07-937-609-19

Query Match 37.2%; Score 42; DB 1; Length 369;
Best Local Similarity 53.3%; Pred. No. 71;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSA 16
DB 70 YFIINLALADLCMAA 84

RESULT 11
US-08-029-170-19
Sequence 19, Application US/08029170
PATENT No. 6169173
GENERAL INFORMATION:
APPLICANT: WANK, Stephen A.

TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-08-029-170-19
Query Match 37.2%; Score 42; DB 2; Length 369;
Best Local Similarity 53.3%; Pred. No. 71;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 YYDITLALAGICGSA 16
Db 70 YFINLALADLCMAA 84
RESULT 12
US-09-443-745-19
Sequence 19, Application US/09443745
Patent No. 6706493
GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: rat substance K receptor
US-09-443-745-19
Query Match 37.2%; Score 42; DB 2; Length 369;
Best Local Similarity 53.3%; Pred. No. 71;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 YYDITLALAGICGSA 16
Db 70 YFINLALADLCMAA 84
RESULT 13
US-09-252-991A-23655
Sequence 23655, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23655
LENGTH: 823
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23655
Query Match 37.2%; Score 42; DB 2; Length 823;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 9 LAGICOSARLVQOLA 23
| | | | | | | | | |
DB 524 LAGICQARLQADAA 538

RESULT 14

US-09-248-796A-27407
Sequence 27407, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 27407
LENGTH: 60
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-27407

Query Match 36.3%; Score 41; DB 2; Length 60;
Best Local Similarity 53.8%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 YDITLALAGIOS 15
| | | | | | | | | |
DB 44 YDINMIVIGICLS 56

RESULT 15

US-09-902-540-10640
Sequence 10640, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 10640
LENGTH: 114
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-10640

Query Match 36.3%; Score 41; DB 2; Length 114;
Best Local Similarity 57.1%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 LALAGICQARLVQ 20
| | | | | | | | | |
DB 12 VWLAGACAVARLIQ 25

Search completed: January 23, 2006, 10:46:50
Job time : 22.0917 secs

THE 1990 PINK (1990)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:53:39 ; Search time 3.21397 Seconds
(without alignments)
72.521 Million cell updates/sec

Title: US-10-501-838a-26
Sequence: 1 MYDDTLALAGICQSARLVQQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA_New:*
1: /cgn2_6/pcdata/2/pubppa/US08_NEW_PUB.pep.*
2: /cgn2_6/pcdata/2/pubppa/US06_NEW_PUB.pep.*
3: /cgn2_6/pcdata/2/pubppa/US07_NEW_PUB.pep.*
4: /cgn2_6/pcdata/2/pubppa/PCT_NEW_PUB.pep.*
5: /cgn2_6/pcdata/2/pubppa/US09_NEW_PUB.pep.*
6: /cgn2_6/pcdata/2/pubppa/US10_NEW_PUB.pep.*
7: /cgn2_6/pcdata/2/pubppa/US11_NEW_PUB.pep.*
8: /cgn2_6/pcdata/2/pubppa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	38.1	42	US-10-512-295A-3	Sequence 3, Appl1
2	43	38.1	185	US-10-453-372-896	Sequence 896, App
3	43	38.1	194	US-10-453-372-894	Sequence 894, App
4	43	38.1	218	US-10-453-372-892	Sequence 892, App
5	41	36.3	305	US-10-520-820-1	Sequence 1, Appl1
6	40.5	35.8	365	US-11-000-463-241	Sequence 241, App
7	40.5	35.8	365	US-11-000-463-713	Sequence 713, App
8	40	35.4	63	US-10-467-657-6638	Sequence 6638, App
9	39	34.5	1015	US-10-957-569-51	Sequence 51, Appl1
10	38	33.6	203	US-10-453-372-890	Sequence 890, App
11	38	33.6	434	US-10-467-657-6894	Sequence 6894, App
12	38	33.6	655	US-10-055-877-40	Sequence 40, Appl1
13	38	33.6	678	US-10-055-877-202	Sequence 202, App
14	38	33.6	687	US-10-055-877-199	Sequence 199, App
15	38	33.6	687	US-10-055-877-201	Sequence 201, App
16	37.5	33.2	231	US-11-082-389-242	Sequence 242, App
17	37	32.7	140	US-10-793-626-2082	Sequence 2082, App
18	37	32.7	367	US-10-821-234-1560	Sequence 1560, App
19	36.5	32.3	481	US-10-467-657-6660	Sequence 4660, App
20	36	31.9	228	US-10-467-657-1242	Sequence 1242, App
21	36	31.9	301	US-10-467-657-7356	Sequence 7356, App
22	36	31.9	491	US-10-763-712A-4	Sequence 4, Appl1
23	36	31.9	919	US-10-858-730-206	Sequence 206, App
24	36	31.9	1268	US-11-052-554A-1	Sequence 1, Appl1
25	35.5	31.4	341	US-10-467-657-154	Sequence 154, App

26	35.5	31.4	341	6	US-10-467-657-5822	Sequence 5822, App
27	35.5	31.4	341	6	US-10-467-657-7476	Sequence 7476, App
28	35.5	31.4	1259	6	US-10-467-657-5510	Sequence 5510, App
29	35	31.0	250	6	US-10-793-626-2310	Sequence 2310, App
30	35	31.0	293	6	US-10-967-671-15	Sequence 15, Appl1
31	35	31.0	305	7	US-11-000-463-391	Sequence 391, App
32	35	31.0	305	7	US-11-000-463-863	Sequence 863, App
33	35	31.0	332	6	US-10-454-437-308	Sequence 308, App
34	35	31.0	345	6	US-10-995-561-975	Sequence 975, App
35	35	31.0	356	6	US-11-194-246-2054	Sequence 2054, App
36	35	31.0	445	7	US-11-194-246-345	Sequence 345, App
37	35	31.0	446	6	US-10-793-626-1836	Sequence 1836, App
38	35	31.0	502	7	US-10-063-703-158	Sequence 158, App
39	35	31.0	502	7	US-11-102-240-158	Sequence 158, App
40	35	31.0	651	7	US-11-124-368A-184	Sequence 184, App
41	34.5	30.5	150	7	US-11-212-443-74	Sequence 74, Appl1
42	34.5	30.5	180	7	US-11-212-443-194	Sequence 194, App
43	34.5	30.5	255	7	US-11-212-443-76	Sequence 76, Appl1
44	34.5	30.5	255	7	US-11-212-443-78	Sequence 78, Appl1
45	34.5	30.5	884	6	US-10-878-556A-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1
US-10-512-295A-3
; Sequence 3, Application US/10512295A
; Publication No. US20050245727A1
GENERAL INFORMATION:
; APPLICANT: Baltzer, Lars
; APPLICANT: Dolphin, Gunnar
; APPLICANT: Liedberg, Bo
; APPLICANT: Lundstrom, Ingemar
; TITLE OF INVENTION: NOVEL POLYPEPTIDE SCAFFOLDS AND USE THEREOF
; FILE REFERENCE: 5848.181USMO
; CURRENT APPLICATION NUMBER: US/10/512,295A
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/SE03/00507
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: SE 0200968-6
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Synthetic Peptide KEX3
US-10-512-295A-3
Query Match 38.1%; Score 43; DB 6; Length 42;
Best Local Similarity 52.9%; Pred. No. 0.52;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 7 LALAGICQSARLVQQLA 23
Db 16 LAARGPCDAQLAEQLA 32
RESULT 2
US-10-453-372-896
; Sequence 896, Application US/10453372
; Publication No. US20060003323A1
GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23

```
;; PRIOR APPLICATION NUMBER: 60/185967
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 09/823187
;; PRIOR FILING DATE: 2001-03-29
;; PRIOR APPLICATION NUMBER: 60/195792
;; PRIOR FILING DATE: 2000-03-10
;; PRIOR APPLICATION NUMBER: 09/839446
;; PRIOR FILING DATE: 2001-03-19
;; PRIOR APPLICATION NUMBER: 60/199476
;; PRIOR FILING DATE: 2000-03-25
;; PRIOR APPLICATION NUMBER: 09/863776
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: 60/208263
;; PRIOR FILING DATE: 2000-05-31
;; PRIOR APPLICATION NUMBER: 09/939398
;; PRIOR FILING DATE: 2001-08-24
;; PRIOR APPLICATION NUMBER: 60/227800
;; PRIOR FILING DATE: 2000-08-25
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1609
;; SOFTWARE: CurSeqList version 0.1
;; SEQ ID NO 896
;; LENGTH: 185
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-453-372-896
```

```
Query Match      38.1%; Score 43; DB 6; Length 185;
Best Local Similarity 52.4%; Pred. No. 2.4;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
Oy      3 YDITLALAGICQSARLVQOLA 23
Db      67 YDSLALDGHIGSARALMVVA 87
```

```
RESULT 3
US-10-453-372-894
;; Sequence 894, Application US/10453372
;; Publication No. US20060003323A1
;; GENERAL INFORMATION:
;; APPLICANT: Alsbrook, et al.
;; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;; FILE REFERENCE: 21402-589 A
;; CURRENT APPLICATION NUMBER: US/10/453,372
;; CURRENT FILING DATE: 2003-06-03
;; PRIOR APPLICATION NUMBER: 09/789390
;; PRIOR FILING DATE: 2001-02-23
;; PRIOR APPLICATION NUMBER: 60/185967
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 09/823187
;; PRIOR FILING DATE: 2001-03-29
;; PRIOR APPLICATION NUMBER: 60/195792
;; PRIOR FILING DATE: 2000-03-10
;; PRIOR APPLICATION NUMBER: 09/839446
;; PRIOR FILING DATE: 2001-03-19
;; PRIOR APPLICATION NUMBER: 60/199476
;; PRIOR FILING DATE: 2000-03-25
;; PRIOR APPLICATION NUMBER: 09/863776
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: 60/208263
;; PRIOR FILING DATE: 2000-05-31
;; PRIOR APPLICATION NUMBER: 09/939398
;; PRIOR FILING DATE: 2001-08-24
;; PRIOR APPLICATION NUMBER: 60/227800
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1609
;; SOFTWARE: CurSeqList version 0.1
;; SEQ ID NO 894
;; LENGTH: 194
;; TYPE: PRT
;; ORGANISM: Homo sapiens
```

```
US-10-453-372-894
```

```
Query Match      38.1%; Score 43; DB 6; Length 194;
Best Local Similarity 52.4%; Pred. No. 2.6;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
Oy      3 YDITLALAGICQSARLVQOLA 23
Db      47 YDSLALDGHIGSARALMVVA 67
```

```
RESULT 4
US-10-453-372-892
;; Sequence 892, Application US/10453372
;; Publication No. US20060003323A1
;; GENERAL INFORMATION:
;; APPLICANT: Alsbrook, et al.
;; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
;; FILE REFERENCE: 21402-589 A
;; CURRENT APPLICATION NUMBER: US/10/453,372
;; CURRENT FILING DATE: 2003-06-03
;; PRIOR APPLICATION NUMBER: 09/789390
;; PRIOR FILING DATE: 2001-02-23
;; PRIOR APPLICATION NUMBER: 60/185967
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 09/823187
;; PRIOR FILING DATE: 2001-03-29
;; PRIOR APPLICATION NUMBER: 60/195792
;; PRIOR FILING DATE: 2000-03-10
;; PRIOR APPLICATION NUMBER: 09/839446
;; PRIOR FILING DATE: 2001-03-19
;; PRIOR APPLICATION NUMBER: 60/199476
;; PRIOR FILING DATE: 2000-03-25
;; PRIOR APPLICATION NUMBER: 09/863776
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: 60/208263
;; PRIOR FILING DATE: 2000-05-31
;; PRIOR APPLICATION NUMBER: 09/939398
;; PRIOR FILING DATE: 2001-08-24
;; PRIOR APPLICATION NUMBER: 60/227800
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1609
;; SOFTWARE: CurSeqList version 0.1
;; SEQ ID NO 892
;; LENGTH: 218
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-453-372-892
```

```
Query Match      38.1%; Score 43; DB 6; Length 218;
Best Local Similarity 52.4%; Pred. No. 2.9;
Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
```

```
Oy      3 YDITLALAGICQSARLVQOLA 23
Db      71 YDSLALDGHIGSARALMVVA 91
```

```
RESULT 5
US-10-520-820-1
;; Sequence 1, Application US/10520820
;; Publication No. US20060003393A1
;; GENERAL INFORMATION:
;; APPLICANT: MUTHAIGIS S.A.
;; TITLE OF INVENTION: Pathogenicity determinants which can be used as targets for devel-
;; TITLE OF INVENTION: means for preventing and controlling bacterial infections and/or
;; FILE REFERENCE: 1621
;; CURRENT APPLICATION NUMBER: US/10/520,820
;; CURRENT FILING DATE: 2005-01-07
;; PRIOR APPLICATION NUMBER: PCT/EP2003/008209
;; PRIOR FILING DATE: 2003-07-09
```

PRIOR APPLICATION NUMBER: FR 0208636
PRIOR FILING DATE: 2002-07-09
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 305
TYPE: PRF
ORGANISM: Escherichia coli
US-10-520-820-1

Query Match 36.3%; Score 41; DB 6; Length 305;
Best Local Similarity 42.9%; Pred. No. 9;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 YDITLALGICGASRLVQOLA 23
DB 69 YDANVDAQIVKASALVTRLA 89

RESULT 6
US-11-000-463-241
Sequence 241, Application US/11000463
Publication No. US2005026423A1
GENERAL INFORMATION:
APPLICANT: Tang, Y Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Chen, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Drmanac, Radcoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIP4CN
CURRENT APPLICATION NUMBER: US/11/000,463
CURRENT FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 241
LENGTH: 365
TYPE: PRF
ORGANISM: Homo sapiens
US-11-000-463-241

Query Match 35.8%; Score 40.5; DB 7; Length 365;
Best Local Similarity 43.5%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 4 DITLALGIC--QSARLVQOLA 23
DB 67 DLMGLAGFCGSMASGHLFKOMA 89

RESULT 7
US-11-000-463-713
Sequence 713, Application US/11000463

Publication No. US2005026423A1
GENERAL INFORMATION:
APPLICANT: Tang, Y Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Chen, Rui-hong
APPLICANT: Qian, Xiaohong B.
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Cao, Yi-Cheng
APPLICANT: Drmanac, Radcoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 785CIP4CN
CURRENT APPLICATION NUMBER: US/11/000,463
CURRENT FILING DATE: 2004-11-29
PRIOR APPLICATION NUMBER: 10/291,265
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 09/922,279
PRIOR FILING DATE: 2001-08-03
PRIOR APPLICATION NUMBER: 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: 09/617,746
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 09/631,451
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 09/633,870
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 944
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 713
LENGTH: 365
TYPE: PRF
ORGANISM: Homo sapiens
US-11-000-463-713

Query Match 35.8%; Score 40.5; DB 7; Length 365;
Best Local Similarity 43.5%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 4 DITLALGIC--QSARLVQOLA 23
DB 67 DLMGLAGFCGSMASGHLFKOMA 89

RESULT 8
US-10-467-657-6638
Sequence 6638, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASTIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 6638
LENGTH: 63
TYPE: PRF
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6638

Query Match 35.4%; Score 40; DB 6; Length 63;

Best Local Similarity 53.3%; Pred. No. 2.5;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 YDITLALAGICQSA 16
Db 46 YNITVASTVACQTA 60

RESULT 9

US-10-957-569-51
; Sequence 51, Application US/10957569
; Publication No. US20050246785A1
; GENERAL INFORMATION:
; APPLICANT: COOK, Zhihong et al.
; TITLE OF INVENTION: PROMOTER, PROMOTER CONTROL ELEMENTS, AND COMBINATIONS, AND USES
; FILE REFERENCE: 2750-1577PUS3
; CURRENT APPLICATION NUMBER: US/10/957,569
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 10/950,321
; PRIOR FILING DATE: 2004-09-23
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-957-569-51

Query Match 34.5%; Score 39; DB 6; Length 1015;
Best Local Similarity 35.0%; Pred. No. 68;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 4 DITLALAGICQSA 23
Db 172 DMTLMIGVCAFLVIGIA 191

RESULT 10

US-10-453-372-890
; Sequence 890, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 890
; LENGTH: 203

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-890

Query Match 33.6%; Score 38; DB 6; Length 203;
Best Local Similarity 47.6%; Pred. No. 19;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 3 YDITLALAGICQSA 23
Db 67 YDSLALDGRPQARALVVA 87

RESULT 11

US-10-467-657-6894
; Sequence 6894, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASTIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 6894
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6894

Query Match 33.6%; Score 38; DB 6; Length 434;
Best Local Similarity 28.6%; Pred. No. 42;
Matches 10; Conservative 6; Mismatches 7; Indels 12; Gaps 1;

Qy 1 MYDITLALAGIC-----QSA 23
Db 283 MDPFIPGVNDGCDRYLCRMEMRQSVRIKQCA 317

RESULT 12

US-10-055-877-40
; Sequence 40, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Rameesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zettnusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Palturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eiesen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir

APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ference
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 655
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-877-40

Query Match 33.6%; Score 38; DB 6; Length 655;
Best Local Similarity 64.3%; Fred. No. 64;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 LAGTOSARLPOOL 22
DB 558 LVGIYVRAQLVQNL 571

RESULT 13
US-10-055-877-202
Sequence 202, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigar, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zerhusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eileen, Andrew
APPLICANT: Wolenc, Adam

APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Vernet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ference
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 202
LENGTH: 678
TYPE: PRT
ORGANISM: Oryctolagus cuniculus
US-10-055-877-202

Query Match 33.6%; Score 38; DB 6; Length 678;
Best Local Similarity 64.3%; Fred. No. 66;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 LAGTOSARLPOOL 22
DB 590 LVGIYVRAQLVQNL 603

RESULT 14
US-10-055-877-199
Sequence 199, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigar, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zerhusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera

```

: APPLICANT: Burgess, Catherine
: APPLICANT: Etsen, Andrew
: APPLICANT: Wolenc, Adam
: APPLICANT: Baumgartner, Jason
: APPLICANT: Shinkets, Richard
: APPLICANT: Gusev, Vladimir
: APPLICANT: Vermet, Corine
: APPLICANT: Taupier Jr., Raymond
: APPLICANT: Pena, Carol
: APPLICANT: Shenoy, Suresh
: APPLICANT: Li, Li
: APPLICANT: Caeman, Stacie
: APPLICANT: Boldog, Ference
: TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
: FILE REFERENCE: 21402-251
: CURRENT APPLICATION NUMBER: US/10/055,877
: PRIOR FILING DATE: 2002-01-22
: PRIOR APPLICATION NUMBER: 60/262,892
: PRIOR FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: 60/263,598
: PRIOR FILING DATE: 2001-01-23
: PRIOR APPLICATION NUMBER: 60/263,799
: PRIOR FILING DATE: 2001-01-24
: PRIOR APPLICATION NUMBER: 60/264,117
: PRIOR FILING DATE: 2001-01-25
: PRIOR APPLICATION NUMBER: 60/264,139
: PRIOR FILING DATE: 2001-01-25
: PRIOR APPLICATION NUMBER: 60/264,478
: PRIOR FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: 60/263,351
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: 60/272,870
: PRIOR FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: 60/275,990
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 60/275,927
: PRIOR FILING DATE: 2001-03-14
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 512
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 199
: LENGTH: 687
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-055-877-199

Query Match          33.6%; Score 38; DB 6; Length 687;
Best Local Similarity 64.3%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy          9 LAGICQSARLVQQL 22
Db          590 LVGIVQRAQLVQAL 603

RESULT 15
US-10-055-877-201
: Sequence 201, Application US/10055877
: Publication No. US20050288241A1
: GENERAL INFORMATION:
: APPLICANT: Decristofaro, Marc
: APPLICANT: Padigar, Muralidhara
: APPLICANT: Miller, Charles
: APPLICANT: Tchernev, Velizar
: APPLICANT: Zhong, Mei
: APPLICANT: Anderson, David
: APPLICANT: Ballinger, Robert
: APPLICANT: Gerlach, Valerie
: APPLICANT: Spyrek, Kimberly
: APPLICANT: Raccelli, Luca
: APPLICANT: Kekuda, Ramesh
: APPLICANT: Guo, Xiaojia
: APPLICANT: Zernuusen, Bryan
```

```

: APPLICANT: Andrew, David
: APPLICANT: Mezes, Peter
: APPLICANT: Patturajan, Meera
: APPLICANT: Burgess, Catherine
: APPLICANT: Etsen, Andrew
: APPLICANT: Wolenc, Adam
: APPLICANT: Baumgartner, Jason
: APPLICANT: Shinkets, Richard
: APPLICANT: Gusev, Vladimir
: APPLICANT: Vermet, Corine
: APPLICANT: Taupier Jr., Raymond
: APPLICANT: Pena, Carol
: APPLICANT: Shenoy, Suresh
: APPLICANT: Li, Li
: APPLICANT: Caeman, Stacie
: APPLICANT: Boldog, Ference
: TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
: FILE REFERENCE: 21402-251
: CURRENT APPLICATION NUMBER: US/10/055,877
: PRIOR FILING DATE: 2002-01-22
: PRIOR APPLICATION NUMBER: 60/262,892
: PRIOR FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: 60/263,598
: PRIOR FILING DATE: 2001-01-23
: PRIOR APPLICATION NUMBER: 60/263,799
: PRIOR FILING DATE: 2001-01-24
: PRIOR APPLICATION NUMBER: 60/264,117
: PRIOR FILING DATE: 2001-01-25
: PRIOR APPLICATION NUMBER: 60/264,139
: PRIOR FILING DATE: 2001-01-25
: PRIOR APPLICATION NUMBER: 60/264,478
: PRIOR FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: 60/263,351
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: 60/272,870
: PRIOR FILING DATE: 2001-03-02
: PRIOR APPLICATION NUMBER: 60/275,990
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 60/275,927
: PRIOR FILING DATE: 2001-03-14
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 512
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 201
: LENGTH: 687
: TYPE: PRT
: ORGANISM: Oryctolagus cuniculus
US-10-055-877-201

Query Match          33.6%; Score 38; DB 6; Length 687;
Best Local Similarity 64.3%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy          9 LAGICQSARLVQQL 22
Db          590 LVGIVQRAQLVQAL 603

Search completed: January 23, 2006, 12:09:42
Job time : 3.21397 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:20:59 ; Search time 76.3319 Seconds
(without alignments)
132.392 Million cell updates/sec

Title: US-10-501-838a-26

Perfect score: 1 MYDITLALGICQSAFLVQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

A_Geneseq_21.*
1: geneeqp1980a.*
2: geneeqp1990a.*
3: geneeqp2000a.*
4: geneeqp2001a.*
5: geneeqp2002a.*
6: geneeqp2003a.*
7: geneeqp2003b.*
8: geneeqp2004a.*
9: geneeqp2005a.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	100.0	23	6	ADBI6909 Escherich
2	113	100.0	23	9	AEB08243 Escherich
3	113	100.0	30	6	ADBI6919 Escherich
4	113	100.0	30	9	AEB08253 Escherich
5	108	95.6	23	6	ADBI6888 Escherich
6	108	95.6	23	9	AEB08220 Escherich
7	108	95.6	24	6	ADBI6908 Escherich
8	108	95.6	24	9	AEB08242 Escherich
9	108	95.6	25	6	ADBI6922 Escherich
10	108	95.6	25	9	ADBI6906 Escherich
11	108	95.6	26	6	ADBI6904 Escherich
12	108	95.6	29	6	ADBI6921 E.coli pe
13	108	95.6	30	6	ADBI6903 Escherich
14	108	95.6	30	9	AEB08239 Escherich
15	108	95.6	31	6	ADBI6918 Escherich
16	108	95.6	31	9	AEB08252 Escherich
17	108	95.6	31	9	AEB08263 Escherich
18	108	95.6	213	9	AEB08263 Escherich
19	104	92.0	217	7	ABO61272 Klebsiell
20	101	89.4	22	6	ADBI6910 Escherich
21	101	89.4	22	9	AEB08244 Escherich
22	101	89.4	29	6	ADBI6920 Escherich
23	101	89.4	29	9	AEB08254 Escherich
24	97	85.8	23	6	ADBI6887 Pasteurel

25	97	85.8	23	9	AEB08219
26	97	85.8	203	9	AEB08262
27	92	81.4	23	6	ADBI6912
28	92	81.4	23	9	AEB08246
29	92	81.4	29	6	ADBI6916
30	92	81.4	29	9	AEB08250
31	91	80.5	209	6	ABM67069
32	89	78.8	23	6	ADBI6886
33	89	78.8	23	9	AEB08218
34	89	78.8	205	4	AAW50230
35	89	78.8	205	9	AEB08261
36	87	77.0	216	7	ADF07708
37	87	77.0	23	6	ADBI6890
38	80	70.8	23	9	AEB08222
39	80	70.8	211	9	AEB08265
40	72	63.7	23	6	ADBI6889
41	72	63.7	23	9	AEB08221
42	72	63.7	204	9	AEB08264
43	48	42.5	23	6	ADBI6896
44	48	42.5	23	9	AEB08228
45	48	42.5	231	3	AAV93280

ALIGNMENTS

RESULT 1
ADBI6909 standard; peptide; 23 AA.
XX
XX
XX
ADBI6909;
XX
XX
20-NOV-2003 (first entry)
XX
XX
Escherichia coli YCFC penetrating peptide 31.
DB
XX
XX
penetrating peptide; epithelial; endothelial; tight junction; diabetes;
XX
XX
infertility; hormone; vitamin deficiency; neurodegenerative;
XX
XX
cardiovascular; haematological; endocrine disorder; obesity;
XX
XX
neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
XX
XX
osteopathic; cytoskeletal; nootropic.
OS
XX
Escherichia coli.
XX
XX
WC02003066859-A2.
XX
XX
14-AUG-2003.
XX
XX
07-FEB-2003; 2003WO-1B000968.
XX
XX
07-FEB-2002; 2002US-0355396P.
XX
XX
(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX
Ben-Sasson SA, Cohen E;
XX
XX
WPI; 2003-697452/66.
XX
XX
New penetrating peptide, useful for preparing a composition for treating
XX
XX
or preventing e.g. endocrine disorders.
XX
XX
Claim 2; Page 15; 60pp; English.
XX
XX
This invention relates to a novel peptide sequences capable of
XX
XX
translocating across a biological barrier. Furthermore, it refers to
XX
XX
methods that use these peptides to facilitate penetration of a
XX
XX
biologically active effector molecule such as a drug or other therapeutic
XX
XX
agent across biological barriers e.g. epithelial or endothelial cells
XX
XX
sealed by tight junctions. This peptide is derived from a bacterial
XX
XX
toxin, an integral membrane or extracellular protein and can comprise an
XX
XX
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX
XX
or enzyme. The effector molecule, however, can comprise for example
XX
XX
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from YCFC of
 CC Becherichia coli and is penetrating peptide 31 of the invention.
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 113; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MYDITTLAAGICQSAARLVQOLA 23
 Db 1 MYDITTLAAGICQSAARLVQOLA 23
 RESULT 2
 AEB08243 ID AEB08243 standard; peptide; 23 AA.
 XX AC AEB08243;
 XX DT 25-AUG-2005 (first entry)
 XX XX Becherichia coli YCFC penetrating peptide 31, SEQ ID NO: 26.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; anlaemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cyostatic;
 KW antinflamatory; hepatotropic; hepatitis B virus infection.
 KW Becherichia coli.
 OS US2005136103-A1.
 XX PN 23-JUN-2005.
 PD 23-JUN-2005.
 XX PD 16-SEP-2004; 2004US-00942300.
 XX PP 17-SEP-2003; 2003US-00664989.
 XX PR 17-SEP-2003; 2003US-00665184.
 XX PR 17-SEP-2003; 2003US-0503615P.
 XX PA (BENS/) BEN-SASSON S A.
 XX PA (COHEN/) COHEN E.
 XX PI Ben-Sasson SA, Cohen E;
 XX DR WPI; 2005-444089/45.
 XX PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 46; SEQ ID NO 26; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transseptal delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Becherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 113; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MYDITTLAAGICQSAARLVQOLA 23
 Db 1 MYDITTLAAGICQSAARLVQOLA 23
 RESULT 3
 ADB16919 ID ADB16919 standard; peptide; 30 AA.
 XX AC ADB16919;
 XX DT 20-NOV-2003 (first entry)
 XX XX Becherichia coli peptide 3 coupled to imaging compound linker, IBW-002V2.
 KW Becherichia coli peptide 3 coupled to imaging compound linker, IBW-002V2.
 KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiac; antiarteriosclerotic; osteopathic; cyostatic; neurotropic;
 KW imaging linker; penetrating peptide; IBW-002V2.
 XX
 XX OS Synthetic.
 XX OS Becherichia coli.
 XX OS
 XX OS
 XX PH Key Location/Qualifiers
 XX FT Modified-site 1
 XX FT Modified-site 30 /note= "N-terminal acetyl"
 XX FT /note= "C-terminal amide"
 XX PN WO2003066859-A2.
 XX PD 14-AUG-2003.
 XX PP 07-FEB-2003; 2003WO-IB000968.
 XX PR 07-FEB-2002; 2002US-0355396P.
 XX

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 PI WPI, 2003-697452/66.
 XX
 DR New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PT
 XX Example 3; Page 40; 60pp; English.
 PS
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalaargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin and
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antihypertensive, osteopathic,
 CC cytoskeletal or neurotrophic activities. This peptide sequence is IBM-002V2,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 XX
 SQ Sequence 30 AA:
 QY
 Query Match 100.0%; Score 113; DB 6; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.8e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 MYDITLALAGICQSARLVQQLA 23
 1 MYDITLALAGICQSARLVQQLA 23
 RESULT 4
 AEB08253
 ID AEB08253 standard; peptide; 30 AA.
 AC AEB08253;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DE Penetrating peptide SEQ ID NO: 36 used in composition for translocation.
 XX
 CC Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 CC antidiabetic; endocrine disease; gastrointestinal disease;
 CC metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 CC degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 CC neurodegenerative disease; neuroprotective; Alzheimers disease;
 CC neurotrophic; neurological disease; parkinsons disease; antiparkinsonian;
 CC dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 CC anticonvulsant; genetic disorder; cardiovascular disease;
 CC cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 CC coronary artery disease; cardiac; vasotrophic; obesity; anorectic;
 CC nutritional disorder; vitamin deficiency; renal disease; nephrotoxic;
 CC genitourinary disease; hematological disease; antiemetic; anemia;
 CC autoimmune disease; immunosuppressive; immune deficiency;
 CC immunostimulant; infectious disease; antimicrobial; infection;
 CC erectile dysfunction; andrology; major depressive disorder;
 CC antidepressant; psychiatric disorder; pain; analgesic;
 CC bacterial infection; antibacterial; viral infection; virostatic;
 CC fungal infection; fungicide; parasitic infection; antiparasitic;
 CC renal failure; antifertility; antirheumatic; cytostatic;
 CC antiinflammatory; hepatotropic; hepatitis B virus infection.

XX Unidentified.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT Misc-difference 1
 FT Misc-difference 1
 FT Misc-difference 30
 FT amino group of lysine is acylated with a fatty acid"
 FT
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHE/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI, 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 63; SEQ ID NO 36; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimers disease,
 CC parkinson's disease, multiple sclerosis, Huntingtons disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, autoimmune diseases, immune deficiencies,
 CC rheumatologic disorders, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is a penetrating peptide (IBM-002V2)
 CC used in the composition for the effective translocation of insulin across
 CC an epithelial barrier using BKC as the counter ion and a rating peptide.
 XX
 SQ Sequence 30 AA:
 QY
 Query Match 100.0%; Score 113; DB 9; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.8e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 MYDITLALAGICQSARLVQQLA 23
 1 MYDITLALAGICQSARLVQQLA 23
 RESULT 5
 ADB16888
 ID ADB16888 standard; peptide; 23 AA.
 XX ADB16888;
 AC

XX 20-NOV-2003 (first entry)
 XX Escherichia coli YCFC penetrating peptide 3.
 DE
 XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 KW intercellity; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; arteriosclerotic;
 KW osteopathic; cytoskeletal; motropic.
 KW Escherichia coli.
 OS
 XX WO2003066859-A2.
 PN
 XX 14-AUG-2003.
 PD
 XX 07-FEB-2003; 2003WO-IB000968.
 PP
 XX 07-FEB-2002; 2002US-0355396P.
 PR
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA
 XX Ben-Sasson SA, Cohen E;
 PI
 XX MPI, 2003-697452/66.
 DR
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PS
 XX Claim 2; Page 14; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC antitoxinant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, arteriosclerotic, osteopathic,
 CC cytoskeletal or motropic activities. This peptide is from YCFC of
 CC Escherichia coli and is penetrating peptide 3 of the invention.
 XX
 XX Sequence 23 AA;
 SQ
 Query Match 95.6%; Score 108; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.4e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YYDITLALAGTQSGARLVQOLA 23
 Db 2 YYDITLALAGTQSGARLVQOLA 23

RESULT 6
 AEB08220
 ID AEB08220 standard; peptide; 23 AA.
 AC AEB08220;
 XX
 XX 25-AUG-2005 (first entry)
 DT
 XX Escherichia coli YCFC penetrating peptide 3, SEQ ID NO: 3.
 DE
 XX

KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW motropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; arteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW gouty/urinary disease; hematological disease; anitaneimic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytoskeletal;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 KW
 XX Escherichia coli.
 OS
 XX US2005136103-A1.
 PN
 XX 23-JUN-2005.
 PD
 XX 16-SEP-2004; 2004US-00942300.
 PP
 XX 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 DR MPI, 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 PS
 XX Claim 53; SEQ ID NO 3; 59pp; English.

CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorder, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.

SQ Sequence 23 AA;

Query Match 95.6%; Score 108; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 23
DB 2 YDITLALAGICQSARLVQOLA 23

RESULT 7
ADBI6908

ID ADBI6908 standard; peptide; 24 AA.

AC ADBI6908;

DT 20-NOV-2003 (first entry)

DE Escherichia coli YCFC penetrating peptide 30.

XX penetrating peptide; epithelial; endocervical; tight junction; diabetes;
XX infertility; hormone; vitamin deficiency; neurodegenerative;
XX cardiovascular; haematological; endocrine disorder; obesity;
XX neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
XX osteopathic; cytoskeletal; neurotropic.

OS Escherichia coli.

PN WO2003066859-A2.

PD 14-AUG-2003.

PF 07-FEB-2003; 2003WO-1B000968.

PR 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PA Ben-Sasson SA, Cohen E;

PI WPI, 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.

PS Claim 2; Page 15; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells
XX sealed by tight junctions. This peptide is derived from a bacterial
XX toxin, an integral membrane or extracellular protein and can comprise an
XX antitoxinant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX or enzyme. The effector molecule, however, can comprise for example
XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
XX factors. The penetrating peptide is useful for the treatment of various
XX conditions including diabetes, infertility, hormone and vitamin
XX deficiencies, neurodegenerative, cardiovascular, haematological and
XX endocrine disorders, as well as obesity and neoplastic disease.
XX Accordingly, the peptides of this invention can be used in compositions
XX that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
XX cytoskeletal or neurotropic activities. This peptide is from YCFC of
XX Escherichia coli and is penetrating peptide 30 of the invention.

XX Sequence 24 AA;

Query Match 95.6%; Score 108; DB 6; Length 24;

Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 23

DB 3 YDITLALAGICQSARLVQOLA 24

RESULT 8
AEB08242

ID AEB08242 standard; peptide; 24 AA.

AC AEB08242;

DT 25-AUG-2005 (first entry)

DE Escherichia coli YCFC penetrating peptide 30, SEQ ID NO: 25.

XX pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
XX antidiabetic; endocrine disease; gastrointestinal disease;
XX metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;
XX neurodegenerative disease; neuroprotective; Alzheimer's disease;
XX neurotropic; neurological disease; parkinson's disease; antiparkinsonian;
XX dementia; multiple sclerosis; immune disorder; Huntington's chorea;
XX cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
XX coronary artery disease; cardiac; vasotrophic; obesity; anorectic;
XX nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
XX genitourinary disease; haematological disease; antineutric; anemia;
XX autoimmune disease; immunosuppressive; immune deficiency;
XX immunostimulant; infectious disease; antimicrobial; infection;
XX erectile dysfunction; andrology; major depressive disorder;
XX antidepressant; psychiatric disorder; pain; analgesic;
XX bacterial infection; antibacterial; viral infection; virocidic;
XX fungal infection; fungicide; parasitic infection; antiparasitic;
XX renal failure; antifertility; antineumatic; cystostatic;
XX antiinflammatory; hepatotropic; hepatitis B virus infection.

OS Escherichia coli.

PN US2005136103-A1.

PD 23-JUN-2005.

PF 16-SEP-2004; 2004US-00942300.

PR 17-SEP-2003; 2003US-00664989.

PR 17-SEP-2003; 2003US-00665184.

PR 17-SEP-2003; 2003US-0503615P.

XX (BENS/) BEN-SASSON S A.

PA (COHE/) COHEN E.

PI Ben-Sasson SA, Cohen E;

PT WPI, 2005-444089/45.

PT Composition used for translocating effectors across barrier such as
PT epithelial cells during treatment of e.g. endocrine disorders comprises
PT effector sequentially coupled with counter ion and hydrophobic agent.

XX Claim 46; SEQ ID NO 25; 59pp; English.

XX The present invention relates to a pharmaceutical composition of
XX penetrating peptides for transmembrane delivery of effector. The
XX invention comprises the effector sequentially coupled with a counter ion
XX and at least one hydrophobic agent, where the effector is selectively
XX encapsulated into a complex. The invention is useful for translocating
XX effectors across a biological barrier such as epithelial cells and
XX endothelial cells during treatment and prevention of disease or
XX pathological conditions (including endocrine disorders, diabetes,
XX infertility, hormone deficiencies, osteoporosis, ophthalmological
XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
XX Parkinson's disease, multiple sclerosis, Huntington's disease,
XX cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
XX coagulable states, coronary disease, cerebrovascular events, metabolic

disorders, obesity, vitamin deficiencies, renal disorders, renal failure, CC
hematological disorders, anemia of different entities, immunologic and CC
rheumatologic disorders, autoimmune diseases, immune deficiencies, CC
infectious diseases, viral infections, bacterial infections, fungal CC
infections, parasitic infections, neoplastic diseases, multi-factorial CC
disorders, impotence, chronic pain, depression, different fibrosis states CC
and short stature) and for mucosal vaccination against anthrax and CC
hepatitis B. The present sequence is the Escherichia coli Y0FC CC
penetrating peptide. This sequence is used in the effective translocation CC
of aminoglycoside antibiotics and antifungal agents across an epithelial CC
barrier. CC
XX

Sequence 24 AA;
SQ

Query Match 95.6%; Score 108; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YYDITLALAGICQSARLVQOLA 23
Db 3 YYDITLALAGICQSARLVQOLA 24

RESULT 9
ADBI6922
ID ADBI6922 standard; peptide; 25 AA.
XX
AC ADBI6922;
XX
DT 20-NOV-2003 (first entry)
XX
DE Escherichia coli peptide 3 linked to recombinant human insulin.
XX
KW recombinant human insulin; epithelial; endothelial; tight junction;
KM diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
KW cardiovascular; haematological; endocrine disorder; obesity;
KM neoplastic disease; neuroprotective; cardiact; antiarteriosclerotic;
KW osteopathic; cyostatic; nootropic; penetrating peptide; diabetes.
XX
OS Synthetic.
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT Modified-site /note= "Penetrating peptide 3"
FT /note= "Recombinant human insulin peptide coupled to the
FT C-terminus"
FT
FN WO2003066859-A2.
XX
PD 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX
XX WPI, 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
XX Example 8; Page 42; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells

sealed by tight junctions. This peptide is derived from a bacterial CC
toxin, an integral membrane or extracellular protein and can comprise an CC
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin CC
or enzyme. The effector molecule, however, can comprise for example CC
insulin, gonadotropin, erythropoietin, granulocyte/macocyte colony CC
stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic CC
factors. The penetrating peptide is useful for the treatment of various CC
conditions including diabetes, infertility, hormone and vitamin CC
deficiencies, neurodegenerative, cardiovascular, haematological and CC
endocrine disorders, as well as obesity and neoplastic disease. CC
Accordingly, the peptides of this invention can be used in compositions CC
that have neuroprotective, cardiact, antiarteriosclerotic, osteopathic, CC
cyostatic or nootropic activities. This peptide sequence consists of the CC
Escherichia coli penetrating peptide 3 coupled to recombinant human CC
insulin in the absence of a detachable linker peptide as a cleavage site,
CC used to deliver insulin across mucosal epithelia as a treatment for
CC diabetes, in an exemplification of the invention. CC
XX

Sequence 25 AA;
SQ

Query Match 95.6%; Score 108; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YYDITLALAGICQSARLVQOLA 23
Db 2 YYDITLALAGICQSARLVQOLA 23

RESULT 10
ADBI6906
ID ADBI6906 standard; peptide; 25 AA.
XX
AC ADBI6906;
XX
DT 20-NOV-2003 (first entry)
XX
DE Escherichia coli peptide 3 linked to linearised insulin receptor peptide.
XX
KW linearised insulin receptor; epithelial; endothelial; tight junction;
KM diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
KW cardiovascular; haematological; endocrine disorder; obesity;
KM neoplastic disease; neuroprotective; cardiact; antiarteriosclerotic;
KW osteopathic; cyostatic; nootropic; penetrating peptide; diabetes.
XX
OS Synthetic.
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT Modified-site /note= "Penetrating peptide 3"
FT /note= "Linearised insulin receptor peptide coupled to
FT the C-terminus"
FT
FN WO2003066859-A2.
XX
PD 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX
XX WPI, 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
XX Example 9; Page 43; 60pp; English.
XX

XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide sequence consists of the
 CC Escherichia coli penetrating peptide 3 coupled to a linearised insulin
 CC receptor, used to deliver insulin across mucosal epithelia as a treatment
 CC for diabetes, in an exemplification of the invention.

SQ Sequence 25 AA;

Query Match 95.6%; Score 108; DB 6; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.6e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDITLALAGICQSARLVQOLA 23
 |||||
 DB 2 YDITLALAGICQSARLVQOLA 23

RESULT 11
 ADB16904
 ID ADB16904 standard; peptide; 26 AA.

AC ADB16904;

DT 20-NOV-2003 (first entry)

DE E_coli penetrating peptide 3 linked to heparin without a cleavage site.

XX heparin; epithelial; endothelial; tight junction; diabetes; infertility;
 KW hormone; vitamin deficiency; neurodegenerative; cardiovascular;
 KW haematological; endocrine disorder; obesity; neoplastic disease;
 KW neuroprotective; cardiac; antiarteriosclerotic; osteopathic;
 KW cyostatic; neurotropic; penetrant peptide.

OS Synthetic.
 OS Escherichia coli.

XX Key Location/Qualifiers

FT Peptide 1..23

FT /note= "Penetrating peptide 3"

FT Modified-site 26

FT /note= "Heparin peptide coupled to the C-terminus"

XX WO2003066859-A2.

XX 14-AUG-2003.

XX PD 07-FEB-2003; 2003WO-IB000968.

XX PR 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX PA Ben-Gabson SA, Cohen E;

XX PI WPI, 2003-697452/66.

XX DR

PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 PS Example 10; Page 44; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide sequence consists of the
 CC Escherichia coli penetrating peptide 3 coupled to heparin in the absence
 CC of a detachable linker peptide as a cleavage site, used to deliver low
 CC molecular weight heparin across mucosal epithelia in an exemplification
 CC of the invention.

SQ Sequence 26 AA;

Query Match 95.6%; Score 108; DB 6; Length 26;
 Best Local Similarity 100.0%; Pred. No. 1.6e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YDITLALAGICQSARLVQOLA 23
 |||||
 DB 2 YDITLALAGICQSARLVQOLA 23

RESULT 12
 ADB16921

ID ADB16921 standard; peptide; 29 AA.

AC ADB16921;

DT 20-NOV-2003 (first entry)

DE E_coli peptide 3 linked to recombinant human insulin via cleavage site.

XX recombinant human insulin; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cyostatic; neurotropic; detachable linker;
 KW penetrating peptide; diabetes.

OS Synthetic.
 OS Escherichia coli.

XX Key Location/Qualifiers

FT Peptide 1..23

FT /note= "Penetrating peptide 3"

FT Cleavage-site 26..29

FT /note= "Detachable linker peptide"

FT Modified-site 29

FT /note= "Recombinant human insulin peptide coupled to the C-terminus"

XX WO2003066859-A2.

XX 14-AUG-2003.

XX PD 07-FEB-2003; 2003WO-IB000968.

XX 07-FEB-2002; 2002US-0355396P.
PR (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX Ben-Sasson SA, Cohen E;
XX WPI; 2003-697452/66.
XX New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
XX Example 8; Page 42; 60pp; English.
XX This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticonagulant, antibiolic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiant, antiatherosclerotic, osteopathic,
CC cyrostatic or neurotropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to recombinant human
CC insulin via the detachable linker peptide as a cleavage site, used to
CC deliver insulin across mucosal epithelia as a treatment for diabetes, in
CC an exemplification of the invention.
XX
XX Sequence 29 AA;
SQ
Query Match 95.6%; Score 108; DB 6; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Gy 2 YYDITLALAGICQSARLVQOLA 23
Db 2 YYDITLALAGICQSARLVQOLA 23
RESULT 13
ADBI6903
ID ADBI6903 standard; peptide; 30 AA.
AC ADBI6903;
XX 20-NOV-2003 (first entry)
DT
XX E_coli penetrating peptide 3 linked to heparin through a cleavage site.
DE
XX heparin; epithelial; endothelial; tight junction; diabetes; infertility;
KW hormone; vitamin deficiency; neurodegenerative; cardiovascular;
KW haematological; endocrine disorder; obesity; neoplastic disease;
KW neuroprotective; cardiant; antiatherosclerotic; osteopathic; cyrostatic;
KW neurotropic; detachable linker; penetrating peptide.
XX
XX Synthetic.
OS Escherichia coli.
XX
XX Key location/Qualifiers
FH Peptide 1..23
FT Cleavage-site /note= "Penetrating peptide 3"
FT 26..29
FT Modified-site /note= "Detachable linker peptide"
FT 30

FT /note= "Heparin peptide coupled to the C-terminus"
XX WO2003066859-A2.
XX 14-AUG-2003.
XX 07-FEB-2003; 2003WO-IB000968.
XX 07-FEB-2002; 2002US-0355396P.
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX Ben-Sasson SA, Cohen E;
XX WPI; 2003-697452/66.
XX New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
XX Example 10; Page 44; 60pp; English.
XX This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticonagulant, antibiolic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiant, antiatherosclerotic, osteopathic,
CC cyrostatic or neurotropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to heparin via the
CC detachable linker peptide as a cleavage site, used to deliver low
CC molecular weight heparin across mucosal epithelia in an exemplification
CC of the invention.
XX
XX Sequence 30 AA;
SQ
Query Match 95.6%; Score 108; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Gy 2 YYDITLALAGICQSARLVQOLA 23
Db 2 YYDITLALAGICQSARLVQOLA 23
RESULT 14
ADBI6905
ID ADBI6905 standard; peptide; 30 AA.
AC ADBI6905;
XX 20-NOV-2003 (first entry)
DT
XX Escherichia coli peptide 3 coupled to imaging compound linker, IBW-002.
DE
XX epithelial; endothelial; tight junction; diabetes; infertility; hormone;
KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
KW cardiant; antiatherosclerotic; osteopathic; cyrostatic; neurotropic;
KW imaging linker; penetrating peptide; IBW-002.
XX
XX Synthetic.
OS Escherichia coli.

XX	Key	Location/Qualifiers
FT	Modified-site	1
FT	Modified-site	/note= "N-terminal acetyl"
FT	Modified-site	30
FT	Modified-site	/note= "C-terminal amide"
XX		
PN	WO2003066859-A2.	
XX		
PD	14-AUG-2003.	
XX		
PP	07-FEB-2003; 2003WO-IB000968.	
XX		
PR	07-FEB-2002; 2002US-0355396P.	
XX		
PA	(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.	
XX		
P1	Ben-Saason SA, Cohen E;	
XX		
DR	WPI, 2003-697452/66.	
XX		
PT	New penetrating peptide, useful for preparing a composition for treating	
PT	or preventing e.g. endocrine disorders.	
XX		
PS	Example 1; Page 33; 60pp; English.	
XX		
CC	This invention relates to a novel peptide sequences capable of	
CC	translocating across a biological barrier. Furthermore, it refers to	
CC	methods that use these peptides to facilitate penetration of a	
CC	biologically active effector molecule such as a drug or other therapeutic	
CC	agent across biological barriers e.g. epithelial or endothelial cells	
CC	sealed by tight junctions. This peptide is derived from a bacterial	
CC	toxin, an integral membrane or extracellular protein and can comprise an	
CC	anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin	
CC	or enzyme. The effector molecule, however, can comprise for example	
CC	inulin, gonadotropin, erythropoietin, granulocyte/monocyte colony	
CC	stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic	
CC	factors. The penetrating peptide is useful for the treatment of various	
CC	conditions including diabetes, infertility, hormone and vitamin	
CC	deficiencies, neurodegenerative, cardiovascular, haematological and	
CC	endocrine disorders, as well as obesity and neoplastic disease.	
CC	Accordingly, the peptides of this invention can be used in compositions	
CC	that have neuroprotective, cardiac, antiatherosclerotic, osteopathic,	
CC	cytostatic or nootropic activities. This peptide sequence is IBM-002,	
CC	the consists of the Escherichia coli penetrating peptide 3 coupled to	
CC	the tmagmg linker peptide used in an exemplification of the invention.	
XX		
5Q	Sequence 30 Aa:	

[illegible]

XX	noctropic; neurological disease; parkinsons disease; antiparkinsonian;
KM	dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
KM	anticonvulsant; genetic disorder; cardiovascular disease;
KV	cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
KV	coronary artery disease; cardiast; vasotropic; obesity; anorectic;
KV	nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
KM	gastrointestinal disease; hematological disease; antanemic; anemia;
KM	autoimmune disease; immunosuppressive; immune deficiency;
KM	immunostimulant; infectious disease; antimicrobial; infection;
KM	erectile dysfunction; andrology; major depressive disorder;
KW	antidepressant; psychiatric disorder; pain; analgesic;
KW	bacterial infection; antibacterial; viral infection; virucide;
KM	fungal infection; fungicide; parasitic infection; antiparasitic;
KM	renal failure; antifertility; antirheumatic; cyrostatic;
KM	antiinflammatory; hepatotropic; hepatitis B virus infection.
XX	
OS	Unidentified.
XX	
PH	Key
FT	Misc-difference 1 Location/Qualifiers
FT	/note= "N-terminal acylated"
FT	Misc-difference 30
FT	/note= "Optionally C-terminal amide, optionally the free
FT	amino group of lysine is acylated with a fatty acid"
XX	
PN	US2005136103-A1.
XX	
PD	23-JUN-2005.
XX	
PP	16-SEP-2004; 2004US-00942300.
XX	
PR	17-SEP-2003; 2003US-00664989.
XX	
PR	17-SEP-2003; 2003US-00665184.
XX	
PR	17-SEP-2003; 2003US-0503615P.
XX	
PA	(BENS/ BEN-SASSON S A.
PA	(COHE/) COHEN E.
PI	Ben-Saeson SA, Cohen E;
XX	
DX	WPI; 2005-444089/45.
XX	
PT	Composition used for translocating effectors across barrier such as
PT	epithelial cells during treatment of e.g. endocrine disorders comprises
PT	effector sequentially coupled with counter ion and hydrophobic agent.
XX	
PS	Claim 63; SEQ ID NO 22; 59pp; English.
XX	
CC	The present invention relates to a pharmaceutical composition of
CC	penetrating peptides for transepithelial delivery of effector. The
CC	invention comprises the effector sequentially coupled with a counter ion
CC	and at least one hydrophobic agent, where the effector is selectively
CC	encapsulated into a complex. The invention is useful for translocating
CC	effectors across a biological barrier such as epithelial cells and
CC	endothelial cells during treatment and prevention of disease or
CC	pathological conditions (including endocrine disorders, diabetes,
CC	infertility, hormone deficiencies, osteoporosis, ophthalmological
CC	disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
CC	Parkinson's disease, multiple sclerosis, Huntington's disease,
CC	cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
CC	coagulable states, coronary disease, cerebrovascular events, metabolic
CC	disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC	hematological disorders, anemia of different entities, immunologic and
CC	rheumatologic disorders, autoimmune diseases, immune deficiencies,
CC	infectious diseases, viral infections, bacterial infections, fungal
CC	infections, parasitic infections, neoplastic diseases, multi-factorial
CC	disorders, impotence, chronic pain, depression, different fibrosis states
CC	and short stature) and for mucosal vaccination against anthrax and
CC	hepatitis B. The present sequence is a penetrating peptide (1BW-002) used
CC	in the composition for mucosal vaccination using a counter ion and a
CC	penetrating peptide.
XX	
XX	Sequence 30 AA;

Query Match 95.6%; Score 108; DB 9; Length 30;
 Best Local Similarity 100.0%; Pred. No. 1.9e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 YYDITLALAGICQSARLVQOLA 23
 |||||
 Db 2 YYDITLALAGICQSARLVQOLA 23

Search completed: January 23, 2006, 10:28:34
 Job time : 77.3319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 11:52:43 ; Search time 52.8297 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-26

Sequence: 1 MYDITLALAGICQARLVQQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PubliSearch Applications_AA_Main:*
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	113	100.0	23	US-10-665-184-26 Sequence 26, Appl
2	113	100.0	23	US-10-942-300-26 Sequence 26, Appl
3	113	100.0	23	US-10-501-838a-26 Sequence 26, Appl
4	113	100.0	30	US-10-665-184-36 Sequence 36, Appl
5	113	100.0	30	US-10-942-300-36 Sequence 36, Appl
6	113	100.0	30	US-10-501-838a-36 Sequence 36, Appl
7	108	95.6	23	US-10-665-184-3 Sequence 3, Appl
8	108	95.6	23	US-10-942-300-3 Sequence 3, Appl
9	108	95.6	23	US-10-501-838a-3 Sequence 3, Appl
10	108	95.6	24	US-10-665-184-25 Sequence 25, Appl
11	108	95.6	24	US-10-942-300-25 Sequence 25, Appl
12	108	95.6	24	US-10-501-838a-25 Sequence 25, Appl
13	108	95.6	25	US-10-501-838a-13 Sequence 23, Appl
14	108	95.6	25	US-10-501-838a-23 Sequence 23, Appl
15	108	95.6	26	US-10-501-838a-21 Sequence 21, Appl
16	108	95.6	29	US-10-501-838a-18 Sequence 18, Appl
17	108	95.6	30	US-10-665-184-22 Sequence 22, Appl
18	108	95.6	30	US-10-942-300-22 Sequence 22, Appl
19	108	95.6	30	US-10-501-838a-20 Sequence 20, Appl
20	108	95.6	30	US-10-501-838a-22 Sequence 22, Appl
21	108	95.6	31	US-10-665-184-35 Sequence 35, Appl
22	108	95.6	31	US-10-942-300-35 Sequence 35, Appl
23	108	95.6	31	US-10-501-838a-35 Sequence 35, Appl
24	108	95.6	213	US-10-665-184-61 Sequence 61, Appl
25	108	95.6	213	US-10-942-300-61 Sequence 61, Appl
26	108	95.6	213	US-10-501-838a-55 Sequence 55, Appl
27	101	89.4	22	US-10-665-184-27 Sequence 27, Appl

28	101	89.4	22	US-10-942-300-27	Sequence 27, Appl
29	101	89.4	22	US-10-501-838a-27	Sequence 27, Appl
30	101	89.4	29	US-10-665-184-37	Sequence 37, Appl
31	101	89.4	29	US-10-942-300-37	Sequence 37, Appl
32	101	89.4	29	US-10-501-838a-37	Sequence 37, Appl
33	97	85.8	23	US-10-665-184-2	Sequence 2, Appl
34	97	85.8	23	US-10-942-300-2	Sequence 2, Appl
35	97	85.8	23	US-10-501-838a-2	Sequence 2, Appl
36	97	85.8	203	US-10-665-184-60	Sequence 60, Appl
37	97	85.8	203	US-10-942-300-60	Sequence 60, Appl
38	97	85.8	203	US-10-501-838a-54	Sequence 54, Appl
39	92	81.4	23	US-10-665-184-29	Sequence 29, Appl
40	92	81.4	23	US-10-942-300-29	Sequence 29, Appl
41	92	81.4	23	US-10-501-838a-29	Sequence 29, Appl
42	92	81.4	29	US-10-665-184-33	Sequence 33, Appl
43	92	81.4	29	US-10-942-300-33	Sequence 33, Appl
44	92	81.4	30	US-10-501-838a-33	Sequence 33, Appl
45	89	78.8	23	US-10-665-184-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-10-665-184-26
; Sequence 26, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent version 3.2
; SEQ ID NO 26
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-665-184-26

Query Match 100.0%; Score 113; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYDITLALAGICQARLVQQLA 23
DB 1 MYDITLALAGICQARLVQQLA 23

RESULT 2
US-10-942-300-26
; Sequence 26, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

PRIOR FILING DATE: 2003-09-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 26
LENGTH: 23
TYPE: PRT
ORGANISM: Escherichia coli
US-10-942-300-26

Query Match 100.0%; Score 113; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDITLALAGICQSARLVQOLA 23
Db 1 MYDITLALAGICQSARLVQOLA 23

RESULT 3

US-10-501-838A-26
Sequence 26, Application US/10501838A
Publication No. US20050215478A1

GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-501 NATL

CURRENT FILING DATE: 2004-07-19

PRIOR APPLICATION NUMBER: PCT/IB03/00968

PRIOR FILING DATE: 2003-02-07

PRIOR APPLICATION NUMBER: US 60/355,396

NUMBER OF SEQ ID NOS: 72

SOFTWARE: PatentIn version 3.2

SEQ ID NO 26
LENGTH: 23
TYPE: PRT
ORGANISM: Escherichia coli
US-10-501-838A-26

Query Match 100.0%; Score 113; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDITLALAGICQSARLVQOLA 23
Db 1 MYDITLALAGICQSARLVQOLA 23

RESULT 4

US-10-665-184-36
Sequence 36, Application US/10665184

Publication No. US20040146549A1

GENERAL INFORMATION:
APPLICANT: Cohen, Elnat

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-501CIP

CURRENT FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: PCT/IB03/00968

PRIOR FILING DATE: 2003-02-07

PRIOR APPLICATION NUMBER: 60/355,396

NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 36
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial

FEATURE:
OTHER INFORMATION: Penetrating peptide
NAME/KEY: MOD_RES
LOCATION: (1)-(1)
OTHER INFORMATION: ACETYLATION
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (30)-(30)
OTHER INFORMATION: wherein Xaa is Lysine-NH2
US-10-665-184-36

Query Match 100.0%; Score 113; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDITLALAGICQSARLVQOLA 23
Db 1 MYDITLALAGICQSARLVQOLA 23

RESULT 5

US-10-942-300-36
Sequence 36, Application US/10942300

Publication No. US20050136103A1

GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel

TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-503

CURRENT FILING DATE: 2004-09-16

PRIOR APPLICATION NUMBER: 10/665,184

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 10/664,989

PRIOR FILING DATE: 2003-09-17

PRIOR APPLICATION NUMBER: 60/503,615

NUMBER OF SEQ ID NOS: 65

SOFTWARE: PatentIn version 3.2

SEQ ID NO 36
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial

OTHER INFORMATION: Acylated Penetrating Peptide

LOCATION: (30)-(30)

OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated

US-10-942-300-36

Query Match 100.0%; Score 113; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDITLALAGICQSARLVQOLA 23
Db 1 MYDITLALAGICQSARLVQOLA 23

US-10-501-838A-36
Sequence 36, Application US/10501838A

Publication No. US20050215478A1

GENERAL INFORMATION:
APPLICANT: Cohen, Elnat

TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a

FILE REFERENCE: 24348-501 NATL

```
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 36
/ LENGTH: 30
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic: penetrating peptide
/ NAME/KEY: MISC FEATURE
/ LOCATION: (27) - (27)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (30) - (30)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ NAME/KEY: MISC FEATURE
/ LOCATION: (30) - (30)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ NAME/KEY: MISC FEATURE
/ LOCATION: (30) - (30)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ OTHER INFORMATION: peptide via the free amino groups of the lysine residue
US-10-501-838A-36
```

```
Query Match          100.0%; Score 113; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MYDITLALAGICQSARLVQQLA 23
Db 1 MYDITLALAGICQSARLVQQLA 23
```

```
RESULT 7
US-10-665-184-3
/ Sequence 3, Application US/10665184
/ Publication No. US20040146549A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-501CIP
/ CURRENT APPLICATION NUMBER: US/10/665,184
/ CURRENT FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
US-10-665-184-3
```

```
Query Match          95.6%; Score 108; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YYDITLALAGICQSARLVQQLA 23
Db 2 YYDITLALAGICQSARLVQQLA 23
```

```
RESULT 8
US-10-942-300-3
/ Sequence 3, Application US/10942300
/ Publication No. US20050136103A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-503
/ CURRENT APPLICATION NUMBER: US/10/942,300
/ CURRENT FILING DATE: 2004-09-16
/ PRIOR APPLICATION NUMBER: 10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 10/664,989
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 60/503,615
/ PRIOR FILING DATE: 2003-09-17
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
US-10-942-300-3
```

```
Query Match          95.6%; Score 108; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YYDITLALAGICQSARLVQQLA 23
Db 2 YYDITLALAGICQSARLVQQLA 23
```

```
RESULT 9
US-10-501-838A-3
/ Sequence 3, Application US/10501838A
/ Publication No. US20050215478A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel A.
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
/ FILE REFERENCE: 24348-501 NATL
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
US-10-501-838A-3
```

```
Query Match          95.6%; Score 108; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YYDITLALAGICQSARLVQQLA 23
Db 2 YYDITLALAGICQSARLVQQLA 23
```

```
RESULT 10
US-10-665-184-25
/ Sequence 25, Application US/10665184
/ Publication No. US20040146549A1
/ GENERAL INFORMATION:
```

```

; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-665-184-25
```

```

Query Match          95.6%; Score 108; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 2 YYDITLALAGICQSARLVQOLA 23
    |||||
DB 3 YYDITLALAGICQSARLVQOLA 24
```

```

RESULT 11
US-10-942-300-25
; Sequence 25, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-942-300-25
```

```

Query Match          95.6%; Score 108; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 2 YYDITLALAGICQSARLVQOLA 23
    |||||
DB 3 YYDITLALAGICQSARLVQOLA 24
```

```

RESULT 12
US-10-501-838A-25
; Sequence 25, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
```

```

; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-501-838A-25
```

```

Query Match          95.6%; Score 108; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 2 YYDITLALAGICQSARLVQOLA 23
    |||||
DB 3 YYDITLALAGICQSARLVQOLA 24
```

```

RESULT 13
US-10-501-838A-19
; Sequence 19, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURES:
; OTHER INFORMATION: Synthetic: penetrating peptide
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (25)-(25)
; OTHER INFORMATION: wherein recombinant human insulin is coupled to the penetrating
; OTHER INFORMATION: peptide via the glycine residue
US-10-501-838A-19
```

```

Query Match          95.6%; Score 108; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 2 YYDITLALAGICQSARLVQOLA 23
    |||||
DB 2 YYDITLALAGICQSARLVQOLA 23
```

```

RESULT 14
US-10-501-838A-23
; Sequence 23, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
```



```

/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 23
/ LENGTH: 25
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic: penetrating peptide
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (25)..(25)
/ OTHER INFORMATION: wherein the penetrating peptide is coupled to a linearized
/ OTHER INFORMATION: insulin receptor, which is in turn coupled to recombinant human
/ OTHER INFORMATION: insulin
US-10-501-838A-23

```

```

Query Match          95.6%; Score 108; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 YYDITLALAGICQSARLVQOLA 23
DB      2 YYDITLALAGICQSARLVQOLA 23

```

```

RESULT 15
US-10-501-838A-21
/ Sequence 21, Application US/10501838A
/ Publication NO: US20050215478A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Saason, Shmuel A.
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
/ TITLE OF INVENTION: Biological Barrier
/ FILE REFERENCE: 24348-501 NATL
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 21
/ LENGTH: 26
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic: penetrating peptide
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (26)..(26)
/ OTHER INFORMATION: wherein heparin is coupled to the penetrating peptide via the
/ OTHER INFORMATION: free amino group of the lysine residue
US-10-501-838A-21

```

```

Query Match          95.6%; Score 108; DB 5; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 YYDITLALAGICQSARLVQOLA 23
DB      2 YYDITLALAGICQSARLVQOLA 23

```

Search completed: January 23, 2006, 12:09:01
 Job time : 53.8297 secs

This page Blank (except for)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 09:32:56 ; Search time 12.0087 Seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-27
Perfect score: 106
Sequence: 1 MYDITLALAGICOSARLVQOLA 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	95.3	208	2 AH0199	conserved hypochet
2	101	95.3	213	2 S19211	ycfc protein - Esc
3	101	95.3	213	2 D90829	hypothetical prote
4	101	95.3	213	2 B85687	hypothetical prote
5	98	92.5	215	2 AB0647	conserved hypochet
6	82	77.4	205	2 I64155	hypothetical prote
7	80	75.5	211	2 D84960	hypothetical prote
8	73	68.9	205	2 B82337	conserved hypochet
9	46.5	43.9	510	2 D96741	hypothetical prote
10	46	43.4	335	2 S70671	lipopolysaccharide
11	45	42.5	188	2 J7C1065	coat protein - bee
12	45	42.5	947	2 AH0891	adenyl-transferase
13	44	41.5	204	2 B82600	conserved hypochet
14	43	40.6	282	2 E71543	hypothetical prote
15	43	40.6	299	2 B81688	conserved hypochet
16	43	40.6	847	2 G95843	conserved hypochet
17	42	39.6	336	2 G71954	lipopolysaccharide
18	42	39.6	340	2 G64554	lipopolysaccharide
19	41.5	39.2	434	2 S77457	sensory transduct
20	41.5	39.2	1215	2 S77457	chitinase A (impor
21	41	38.7	274	2 S76154	hypothetical prote
22	41	38.7	317	2 AF0973	lipopolysaccharide
23	41	38.7	317	2 A44156	lipopolysaccharide
24	41	38.7	319	2 A40619	lipopolysaccharide
25	41	38.7	330	2 D86038	heptosyl transfera
26	41	38.7	330	2 C91191	ribx protein - Shi
27	41	38.7	395	2 S14963	conserved hypochet
28	41	38.7	455	2 F82345	protein F53H1.1 [1
29	41	38.7	646	2 H88637	

30	41	38.7	1273	2 AP2421	hypothetical prote
31	40.5	38.2	1800	2 A11918	serine/threonine k
32	40	37.7	150	2 B84034	transcription regu
33	40	37.7	177	2 AB3269	outer membrane lip
34	40	37.7	260	2 A95927	probable transcrip
35	40	37.7	295	2 D82443	transcription regu
36	40	37.7	405	2 A25379	saccharopepsin (EC
37	40	37.7	411	2 T00567	hypothetical prote
38	40	37.7	449	2 B83456	probable transport
39	40	37.7	457	2 A95940	probable integral
40	40	37.7	514	2 B96653	hypothetical prote
41	40	37.7	531	2 C95338	hypothetical prote
42	40	37.7	538	2 A41042	microbial metallopro
43	40	37.7	551	2 A12964	hypothetical prote
44	40	37.7	690	2 D98318	ABC protein AGR_L
45	40	37.7	759	2 D70422	cellulose synthase

ALIGNMENTS

RESULT 1
AH0199
conserved hypothetical protein YP01637 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #ext_change 09-Jul-2004
C/Accession: AH0199
R/Parthill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
H. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MWID:21470413; PMID:11586360
A/Accession: AH0199
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-208 <KUR>
A/Cross-references: UNIPROT:Q8ZF06; UNIPARC:UPI00000CD828; GB:AL590842; PIDD:CA090459.1
C/Genetics:
A/Gene: YP01637

Query Match 95.3% Score 101; DB 2; Length 208;
Best Local Similarity 100.0% Pred. No. 3.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICOSARLVQOLA 22
DB 6 YDITLALAGICOSARLVQOLA 26

RESULT 2
S19211
ycfc protein - Escherichia coli (strain K-12)
C/Species: Escherichia coli
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #ext_change 09-Jul-2004
C/Accession: S19211; A64858
R/Green, S.M.; Drabbe, W.T.
submitted to the EMBL Data Library, May 1991
A/Description: Molecular analysis of the purB-phoP region of Escherichia coli K12.
A/Reference number: S19210
A/Accession: S19211
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-213 <GRB>
A/Cross-references: UNIPROT:P25746; UNIPARC:UPI000013A673; EMBL:X59307; NID:942582; PIDD:
R/Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shaq, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MWID:97426617; PMID:9278503
A/Accession: A64858
A/Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-213 <BLAT>
A:Cross-references: UNIPARC:UPI000013A673; GB:AE000213; GB:U00096; NID:G1787371; PIDN:AA
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ynfC
A:start codon: GTG

Query Match 95.3%; Score 101; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 22
DB 6 YDITLALAGICQSARLVQOLA 26

RESULT 3
D90829
hypothetical protein Ecs1604 [imported] - Escherichia coli (strain O157:H7, substrain RD
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D90829
R:Havasi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gawwara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A95629; MUID:21156231; PMID:11258796
A:Accession: D90829
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <HAY>
A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:BA000007; PIDN:BA035027.1;
A:Experimental source: strain O157:H7, substrain RMD 050952
C:Genetics:
A:Gene: Ecs1604

Query Match 95.3%; Score 101; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 22
DB 6 YDITLALAGICQSARLVQOLA 26

RESULT 4
B85687
hypothetical protein ynfC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85687
R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamoules, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: UNIPROT:O8X736; UNIPARC:UPI00000D0AD2; GB:AE005174; NID:G12514786; F
C:Genetics:
A:Gene: ynfC

Query Match 95.3%; Score 101; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 22
DB 6 YDITLALAGICQSARLVQOLA 26

RESULT 5
AB0647
conserved hypothetical protein STY1273 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typh
A>Note: this species has also been called Salmonella typh
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0647
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0647
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-215 <PAR>
A:Cross-references: UNIPARC:UPI0000059F77; GB:AUS13382; PIDN:CAD08357.1; PID:G16502402;
C:Genetics:
A:Gene: STY1273

Query Match 92.5%; Score 98; DB 2; Length 215;
Best Local Similarity 95.2%; Pred. No. 1.1e-08;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 22
DB 8 YDITLALAGICQSARLVQOLA 28

RESULT 6
164155
hypothetical protein H10638 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: 164155
R:Flitschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkease, E.F.; Kerlavage, A
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodex, A.; Kelley, J.M.; Weidman, J
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: 164155
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-205 <TIGR>
A:Cross-references: UNIPROT:P44796; UNIPARC:UPI000013A674; GB:U32747; GB:L42023; NID:G15
A>Note: best homolog was a hypothetical protein from Escherichia coli

Query Match 77.4%; Score 82; DB 2; Length 205;
Best Local Similarity 76.2%; Pred. No. 5e-06;
Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 22
DB 5 HDIVLALAGVCQSAKLVHQLA 25

RESULT 7
D84960
hypothetical protein [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: D84960
R:Shigenobu, S.; Matsumoto, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A:Reference number: A84930; MUID:20445173; PMID:10993077
A:Accession: D84960

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-211 <STO>
A/Cross-references: UNIPARC:UPI000005BAPF; GB:AP00398; GSPDB:GN00144
A/Experimental source: strain APS
C/Genetics:
A/Gene: ycfC; BU262

Query Match 75.5%; Score 80; DB 2; Length 211;
Best Local Similarity 89.5%; Pred. No. 1.1e-05;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ITTLALAGICQSARLVQOLA 22
|||:|||||:|||||:
DB 7 ITTLALAGICQSARLVQOLA 25

RESULT 8
B82237
conserved hypothetical protein VC1127 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: B82237
R/Author: J. E. Eisen, J. A. Nelson, W. C. Clayton, R. A. Gwin, M. L. Dodson, R. J. Chardon, D. J. Ermolaeva, M. D. J. Vamathevan, J. Baas, S. Qin, H. Dragol, I. Sellers, F. L. R. R. Mekalanos, J. J. Venter, J. C. Fraser, C. M.
N/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: B82237
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-205 <HEI>
A/Cross-references: UNIPROT:Q9KSY9; UNIPARC:UPI00000C2B84; GB:AE004193; GB:AE003852; NID:
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A/Gene: VC1127
A/Map position: 1

Query Match 68.9%; Score 73; DB 2; Length 205;
Best Local Similarity 63.6%; Pred. No. 0.00016;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MYDITLALAGICQSARLVQOLA 22
|||:|||||:|||||:
DB 5 YDRTIAFAGICQAVLVQOVA 26

RESULT 9
D96741
hypothetical protein P17M19.2 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: D96741
R/Author: A. Ecker, J. R. Palm, C. J. Federspiel, N. A. Kaul, S. White, O. Alonso, C. W. Chung, M. K. Conn, L. Conway, A. B. Conway, A. R. Creasy, T. H. Dewar, K. J. Jensen, N. P. Hughes, B. J. Hultzer, L.
N/Title: 408, 816-820, 2000
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: D96741
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-510 <STO>
A/Cross-references: UNIPROT:Q9LE20; UNIPARC:UPI000009FSD8; GB:AE005173; NID:G6978912; P:
C/Genetics:
A/Gene: P17M19.2
A/Map position: 1

Query Match 43.9%; Score 46.5; DB 2; Length 510;
Best Local Similarity 52.4%; Pred. No. 9.3;
Matches 11; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 1 MYDITLALAGICQSARLVQOL 21
|||:|||||:|||||:
DB 314 MYTVPMALAG-CVSARVGNEL 333

RESULT 10
S70671
lipopolysaccharide heptosyltransferase (EC 2.4.99.-) I rfac [similarity] - Bordetella p.
C/Species: Bordetella pertussis
C/Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C/Accession: S70671
R/Author: A. J. Maekell, D.
N/Title: Microbiol. 19, 37-52, 1996
A/Reference number: S70671; MUID:96419162; PMID:8821935
A/Accession: S70671
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-335 <ALT>
A/Cross-references: UNIPROT:Q45375; UNIPARC:UPI00000A9310; EMBL:X90711; NID:992967; P:
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
C/Genetics:
A/Gene: rfac
C/Superfamily: heptosyltransferase I
C/Keywords: glycosyltransferase

Query Match 43.4%; Score 46; DB 2; Length 335;
Best Local Similarity 47.4%; Pred. No. 7.6;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQO 20
|||:|||||:|||||:
DB 89 YDVLDMQGLKSARLVQO 107

RESULT 11
JC1065
coat protein - beet necrotic yellow vein mosaic virus
C/Species: beet necrotic yellow vein mosaic virus
C/Date: 02-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C/Accession: JC1065
R/Author: H. J. Liu, Y. Yu, J. L. Cai, Z. N. Ao, G. M. Yang, L. L.
N/Title: Chinese J. Biotechnol. 9, 147-151, 1993
A/Reference number: JC1065
A/Accession: JC1065
A/Molecule type: mRNA
A/Residues: 1-188 <YAO>
A/Cross-references: UNIPROT:Q86757; UNIPARC:UPI00000P0F13
C/Keywords: coat protein

Query Match 42.5%; Score 45; DB 2; Length 188;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARL 17
|||:|||||:|||||:
DB 143 YTLTVSLAGLAQALRL 158

RESULT 12
AH0891
adenyl-transferase [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C/Species: Salmonella enterica subsp. enterica serovar Typhimurium
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AH0891
R/Author: J. J. Dougan, G. J. James, K. D. J. Thomson, N. R. J. Pickard, D. J. Wain, J. J. Churcher

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:31:19 ; Search time 70.131 Seconds
(without alignments)

221.323 Million cell updates/sec

Title: US-10-501-838a-27
Perfect score: 106
Sequence: 1 MYDITLALGICQSARLVQQLA 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	95.3	208	1 Y1637 YERPE	O821g6 yersinia pe
2	101	95.3	208	1 O669Q3 YERPS	O669g3 yersinia ps
3	101	95.3	213	1 YCFC BC057	O8x736 escherichia
4	101	95.3	213	1 YCFC BC057	O8f1b7 escherichia
5	101	95.3	213	1 YCFC BC057	P257f6 escherichia
6	101	95.3	213	1 YCFC BC057	O831f8 shigella fl
7	98	92.5	213	1 YCFC BC057	O8z7h0 salmonella
8	98	92.5	213	1 YCFC BC057	O5pmj3 salmonella
9	98	92.5	213	1 YCFC BC057	O57qcl salmonella
10	94	88.7	215	2 O57QCL SALPA	O55v55 manheimia
11	93	87.7	204	2 O65VU5 MANSW	O6d4e8 erwinia car
12	92	86.8	203	2 O6DAE8 ERWCT	O9c3j8 pasteurella
13	90	84.9	202	1 Y1850 PASMU	O7n3p4 photorhabdu
14	89	84.0	208	1 Y2805 PHOOL	P47366 haemophilus
15	82	77.4	205	1 Y638 HARTN	O4qms9 haemophilus
16	82	77.4	205	2 O4QMS9 HARTN	P57350 buchnera ap
17	80	75.5	211	1 Y262_BUCAT	O7v135 haemophilus
18	75	70.8	215	1 Y1650_HABDU	O9kx99 vibrio para
19	74	68.9	205	1 Y1129_VIBPA	O7mlt5 vibrio vuln
20	73	68.9	205	1 Y1129_VIBPA	O8d8r4 vibrio vuln
21	72	67.9	205	1 Y1342_VIBVU	O6e3e6 vibrio fusc
22	72	67.9	205	2 O5E3W6_VIBF1	O9sana buchnera ap
23	71	67.0	205	2 Y243_BUCBP	O61t19 photobacter
24	65	61.3	217	1 O6LTI9 PHOPR	O9e8x7 rhizobium 1
25	64	60.4	205	2 O98ER3 RHILC	O6oc88 methylococc
26	48	45.3	197	2 O6OCAR8 METCA	O726u4 desulfovibr
27	48	45.3	207	2 O6Q122_DBSVH	O6q122 rectus nov
28	48	45.3	599	2 O6Q122_DBSVH	O6q122 rectus nov
29	48	45.3	925	2 O6Q122_DBSVH	O6q122 rectus nov
30	48	45.3	3223	2 O6Q122_DBSVH	O6q122 rectus nov
31	48	45.3	3223	2 O6Q122_DBSVH	O6q122 rectus nov

32	48	45.3	3223	2 O8VNR6_ECOLI	O8vnr6 escherichia
33	48	45.3	3223	2 O8VNR6_ECOLI	O8vnr6 escherichia
34	48	45.3	3223	2 O8VNR6_ECOLI	O8vnr6 escherichia
35	48	45.3	3223	2 O8VNR6_ECOLI	O8vnr6 escherichia
36	48	45.3	3223	2 O8VNR6_ECOLI	O8vnr6 escherichia
37	47	44.3	3208	2 O8VNR6_ECOLI	O8vnr6 escherichia
38	47	44.3	3208	2 O8VNR6_ECOLI	O8vnr6 escherichia
39	46.5	43.9	510	2 O91E20_ARATH	O91e20 arabidopsis
40	46	43.4	266	2 O7S4H4_NEUCR	O7s4h4 neuropept
41	46	43.4	320	2 O7W235_BORPA	O7w235 bordetella
42	46	43.4	320	2 O7W235_BORPA	O7w235 bordetella
43	46	43.4	335	2 O88002_BORBR	O88002 bordetella
44	46	43.4	335	2 O88002_BORBR	O88002 bordetella
45	46	43.4	473	2 O5V518_HALMA	O5v518 haloarcula

ALIGNMENTS

RESULT 1
Y1637 YERPE STANDARD; PRT; 208 AA.
AC O8ZP05; O8D0Q1;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein YP01637/Y1798/YP1767.
GN OrderedLocustNames=YPO1637, Y1798, YP1767,
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
NCBI_TaxID=632;
RX [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=CO-92 / Biovar Orientalis;
RC MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RX [2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=KIM5 / Biovar Mediaevalis;
RC MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liang P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.B., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM5";
RJ J. Bacteriol. 184:4601-4611 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cai B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
avirulent to humans";
RJ DNA Res. 11:1179-1197 (2004).
CC -!- SIMILARITY: Belongs to the YP0274 family.

CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its

NS Name=yjcF; OrderedLocuNames=21861, ECs1604;
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21079935; PubMed=1120655; DOI=10.1038/35054089;
 RA Perina N.T., Plunkett G., III, Burdand V., Mau B., Glaeser J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postaf G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobbeck E.J., Davis N.W., Lam A., Dimalanta B.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7,"
 RL Nature 409:529-533(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Tida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kudara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AE005174; AAC55958.1; -; Genomic DNA.
 DR EMBL; BA000007; BAB35027.1; -; Genomic DNA.
 DR PIR; B85687; B85687.
 DR PIR; D90829; D90829.
 DR SMR; O8X736; 2-213.
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 213 AA; 22947 MW; E26E9678C3844E2 CRC64;
 QY Query Match 95.3%; Score 101; DB 1; Length 213;
 DB Best Local Similarity 100.0%; Pred. No. 2e-08;
 DB Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YDITLALAGTQSAKRLVQOLA 22
 DB 6 YDITLALAGTQSAKRLVQOLA 26
 RESULT 4
 YCFC_ECOL6 STANDARD; PRT; 213 AA.
 AC Q8RI17;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein yjcF.
 GN Name=yjcF; OrderedLocuNames=c1511;
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=06:1H1 / CFT073 / ATCC 700928 / UPEC;
 RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Raso D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic *Escherichia coli*."
 CC Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL: AB016759; AAN79980.1; ALT_INIT; Genomic_DNA.
 DR SMR: Q8F1B7; 2-213.
 DR HAMAP: MF_00695; -; 1.
 DR InterPro: IPR007451; DUF489.
 DR Pfam: PF04356; DUF489; 1.
 DR KEGG: Complete proteome; Hypothetical protein.
 DR SEQUENCE 213 AA; 22934 MW; 71C8F96F8EA37BB CRC64;
 SO

Query Match 95.3%; Score 101; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICGASRLVQOLA 22
 DB 6 YDITLALAGICGASRLVQOLA 26

RESULT 5
 YCFC_ECOLI STANDARD; PRT; 213 AA.
 ID YCFC_ECOLI
 AC P25746;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC (ORF-23).
 GN Name=ycfC; OrderedLocustNames=01132;
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OC NCBI_TaxID=562;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K12;
 RL Green S.M., Drabble W.T.;
 RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K12;
 RC MEDLINE=92104952; PubMed=1729205;
 RA He B., Smith J.M., Zalkin H.;
 RT "Escherichia coli purB gene: cloning, nucleotide sequence, and
 RT regulation by purR."
 RL J. Bacteriol. 174:130-136(1992).
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12 / MG1655;
 RC MEDLINE=9742617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12."
 RL Science 277:1453-1474(1997).
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12;
 RC MEDLINE=97061202; PubMed=8905332;
 RX Oshima T., Alba H., Bada T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horichi T.;
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map."
 RL DNA Res. 3:137-155(1996).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to
 CC frameshifts.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL: X59307; CAA41995.1; -; Genomic_DNA.
 DR EMBL: W74924; -; NOT_ANNOTATED_CDS; Genomic_DNA.
 DR EMBL: U00096; AAC74216.1; -; Genomic_DNA.
 DR EMBL: D90748; BAA35954.1; -; Genomic_DNA.
 DR EMBL: D90749; BAA35963.1; -; Genomic_DNA.
 DR PIR: S19211; S19211.
 DR PDB: 1QZ4; X-ray; A=2-213.
 DR PDB: 1SD1; X-ray; A=2-213.
 DR ECHOBASE; EB1321; -.
 DR RCOGENE; EG11345; ycfC.
 DR HAMAP: MF_00695; -; 1.
 DR InterPro: IPR007451; DUF489.
 DR Pfam: PF04356; DUF489; 1.
 DR 3D-structure; Complete proteome; Hypothetical protein.
 DR SEQUENCE 213 AA; 22948 MW; E26E9698C3E542 CRC64;
 SO

Query Match 95.3%; Score 101; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICGASRLVQOLA 22
 DB 6 YDITLALAGICGASRLVQOLA 26

RESULT 6
 YCFC_SHIFL STANDARD; PRT; 213 AA.
 ID YCFC_SHIFL
 AC Q81B8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC.
 GN Name=ycfC; OrderedLocustNames=Sf1151, S1234;
 OS *Shigella flexneri*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Shigella*.
 OC NCBI_TaxID=623;
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=301 / Serotype 2a;
 RC MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RX Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
 RT through comparison with genomes of *Escherichia coli* K12 and O157."
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RC MEDLINE=22590274; PubMed=12704152;
 RX DOI=10.1128/JAI.71.5.2775-2786.2003;

RA Wel J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Kunyen-Jenecky L.U., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of *Shigella*
 RT *flexneri* serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: AE005674; AAN42768.1; ALT_INIT; Genomic_DNA.
 CC EMBL: AB016982; AAP1657.1; -; Genomic_DNA.
 CC SMR: Q83LF8; 2-213.
 CC HAMAP: MF_00695; -; 1.
 CC InterPro: IPR007451; DUF489.
 CC Pfam: PF04356; DUF489; 1.
 CC Complete proteome; Hypothetical protein.
 CC SEQUENCE 213 AA; 22894 MW; 0B0BEC698C29FB4D CRC64;

Query Match 95.3%; Score 101; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 22
 |||||:|||||
 DB 6 YDITLALAGICQSARLVQOLA 26

RESULT 7
 YCFC_SALTY STANDARD; PRT; 213 AA.
 AC Q827H0;
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC.
 GN Name=ycfC; OrderedLocustNames=STY1273, t1687;
 OS Salmonella typhi.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Salmonella.
 CC NCBI_TaxId=601;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowling L., White N., Farrer J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jarvis K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*
 RT *enterica* serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 CC [2]
 CC NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RX DOI=10.1128/JB.185.7.2330-2337.2003;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: AL637269; CAD08357.1; ALT_INIT; Genomic DNA.
 CC EMBL: AE016839; AAO69312.1; ALT_INIT; Genomic_DNA.
 CC SMR: Q827H0; 2-213.
 CC HAMAP: MF_00695; -; 1.
 CC InterPro: IPR007451; DUF489.
 CC Pfam: PF04356; DUF489; 1.
 CC Complete proteome; Hypothetical protein.
 CC SEQUENCE 213 AA; 22942 MW; 63192D1C80665D58 CRC64;

Query Match 92.5%; Score 98; DB 1; Length 213;
 Best Local Similarity 95.2%; Pred. No. 6.4e-08;
 Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 22
 |||||:|||||
 DB 6 YDITLALAGICQSARLVQOLA 26

RESULT 8
 YCFC_SALTY STANDARD; PRT; 213 AA.
 AC Q82P25;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC.
 GN Name=ycfC; OrderedLocustNames=STM1233;
 OS Salmonella typhimurium.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Salmonella.
 CC NCBI_TaxId=602;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 CC -1- SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: AE008754; AAL20162.1; ALT_INIT; Genomic_DNA.
 CC SMR: Q82P25; 2-213.
 CC STyGene: SG72727; ycfC.
 CC HAMAP: MF_00695; -; 1.
 CC InterPro: IPR007451; DUF489.
 CC Pfam: PF04356; DUF489; 1.
 CC Complete proteome; Hypothetical protein.
 CC SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match 92.5%; Score 98; DB 1; Length 213;
 Best Local Similarity 95.2%; Pred. No. 6.4e-08;
 Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSARLVQOLA 22
 |||||:|||||
 DB 6 YDITLALAGICQSARLVQOLA 26

RESULT 9

Q5PMJ3_SALPA PRELIMINARY; PRT; 213 AA.

ID Q5PMJ3;

AC 0570C1;

DT 01-FEB-2005 (TREMBlrel. 29, Created)

DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)

DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)

DE Hypothetical protein yf6c.

GN Name=yf6c; OrderedLocNames=SPAL617;

OS Salmonella paratyphi-a.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OC NCB1_TaxID=54388;

OX NCB1_TaxID=54388;

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 9150;

RX PubMed=1531882; DOI=10.1038/ng1470;

RA McClelland M., Sanderson K.R., Clifton S.W., Lacroix P.,

RA Porrolik S., Sabo A., Meyer R., Bierl T., Ozereky P., McCallan M.,

RA Hartline C.R., Wang C., Nguyen C., Bergoff A., Elliott G.,

RA Kohberg S., Strong C., Du F., Carter J., Kremicki C., Layman D.,

RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Mink P.,

RA Delaunay K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,

RA Splich J., Wilson R.K.;

RT "Comparison of genome degradation in Paratyphi A and Typhi, human-

retracted serovars of Salmonella enterica that cause typhoid.";

RL Nat. Genet. 36:1268-1274(2004).

DR EMBL; CP000026; AAV7544.1; -; Genomic_DNA.

DR SMR; Q5PMJ3; 2-213.

DR InterPro; IPR007451; DUF489.

DR Pfam; PF04356; DUF489; 1.

KW Complete proteome; Hypothetical protein.

SC SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match 92.5%; Score 98; DB 2; Length 213;

Best Local Similarity 95.2%; Pred. No. 6.4e-08;

Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSAKRVQOLA 22

DB 6 YDITLALAGICQSAKRVQOLA 26

RESULT 10

Q570C1_SALCH PRELIMINARY; PRT; 215 AA.

ID Q570C1;

AC 0570C1;

DT 10-MAY-2005 (TREMBlrel. 30, Created)

DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)

DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)

DE Hypothetical protein yf6c.

GN Name=yf6c; OrderedLocNames=SC1184;

OS Salmonella choleraesuis (Salmonella enterica).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OC NCB1_TaxID=591;

OX NCB1_TaxID=591;

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=SC-867;

RX PubMed=15781495;

RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,

RA Wang H.-B., Lee Y.-S.;

RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a

highly invasive and resistant zoonotic pathogen.";

RL Nucleic Acids Res. 33:1690-1698(2005).

DR EMBL; AE017220; AAK65090.1; -; Genomic DNA.

KW Complete proteome; Hypothetical protein.

SC SEQUENCE 215 AA; 23157 MW; C153820629FA8BD CRC64;

Query Match 88.7%; Score 94; DB 2; Length 215;

Best Local Similarity 95.0%; Pred. No. 3.1e-07;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSAKRVQOLA 21

DB 8 YDITLALAGICQSAKRVQOLA 27

RESULT 11

Q65VW5_MANSN PRELIMINARY; PRT; 204 AA.

ID Q65VW5;

AC 065VW5;

DT 25-OCT-2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)

DE Hypothetical protein.

GN OrderedLocNames=MS0298;

OS Mannheimia succiniciproducens (strain MBE155E).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Mannheimia.

OC NCB1_TaxID=221988;

OX NCB1_TaxID=221988;

RP NUCLEOTIDE SEQUENCE.

RC PubMed=15378067; DOI=10.1038/nbt1010;

RX Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,

RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;

RA "The genome sequence of the capnophilic rumen bacterium Mannheimia

succiniciproducens.";

RL Nat. Biotechnol. 22:1275-1281(2004).

DR EMBL; AE016827; AAU36905.1; -; Genomic_DNA.

DR InterPro; IPR007451; DUF489.

DR Pfam; PF04356; DUF489; 1.

KW Complete proteome; Hypothetical protein.

SC SEQUENCE 204 AA; 23036 MW; 3D64A70A30C3CCB CRC64;

Query Match 87.7%; Score 93; DB 2; Length 204;

Best Local Similarity 85.7%; Pred. No. 4.3e-07;

Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YDITLALAGICQSAKRVQOLA 22

DB 6 YDITLALAGICQSAKRVQOLA 26

RESULT 12

Q6D4B8_ERWCT PRELIMINARY; PRT; 212 AA.

ID Q6D4B8;

AC 06D4B8;

DT 25-OCT-2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)

DE Hypothetical protein.

GN OrderedLocNames=SCA2443;

OS Erwilia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Pectobacterium.

OC NCB1_TaxID=29471;

OX NCB1_TaxID=29471;

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RC STRAIN=SCRI 1043 / ATCC BAA-672;

RX PubMed=15263089; DOI=10.1073/pnas.0402424101;

RA Beil K.S., Sebahia M., Fritchard L., Holden M.T.G., Hyman L.J.,

RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,

RA Alkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,

RA Fraser A., Hance Z., Hauser H., Jagsle K., Moule S., Norbertczak H.,

RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,

RA Salmon G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;

RT "Genome sequence of the enterobacterial phytopathogen Erwilia

carotovora subsp. atroseptica and characterization of virulence

factors.";

RT Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).

DR EMBL; BX950851; CAG75345.1; -; Genomic_DNA.

DR SMR; Q6D4B8; 2-207.

DR InterPro; IPR007451; DUF489.

Query Match 87.7%; Score 93; DB 2; Length 204;

Best Local Similarity 85.7%; Pred. No. 4.3e-07;

DR Pfam: PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 212 AA; 22892 MW; A67813005F32B150 CRC64;
Query Match 86.8%; Score 92; DB 2; Length 212;
Best Local Similarity 90.5%; Pred. No. 6.5e-07;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 YDITLALAGICQSARLVQOLA 22
Db 6 YETTLALAGICQSARLVQOLA 26
RESULT 13
Y1850.PASMU
ID Y1850.PASMU STANDARD; PRT; 203 AA.
AC Q9CJY8;
DT 10-OCT-2003 (Rel. 42, Created)
DR 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein PM1850.
GN OrderedLocustNames=PM1850;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; AE006223; AK03934.1; -; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 203 AA; 22717 MW; 5EB91EBF0E181EDF CRC64;
Query Match 84.9%; Score 90; DB 1; Length 203;
Best Local Similarity 81.0%; Pred. No. 1.4e-06;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 YDITLALAGICQSARLVQOLA 22
Db 5 YDITLALAGICQSARLVQOLA 25
RESULT 14
Y2805.PHOLL
ID Y2805.PHOLL STANDARD; PRT; 208 AA.
AC Q7N3B4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein plu2805.
GN OrderedLocustNames=plu2805;
OS Photobacterium luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photobacterium.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TT01;

RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangoul L., Buchrieser C., Glavaud A.,
RA Taourit S., Bocs S., Bouraux-Eude C., Chandler M., Charles J.-F.,
RA Daaba E., Derose R., Berzelle S., Freysinet G., Gaudinault S.,
RA Medigne C., Ianols A., Powell K., Sigulier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunet F.;
RT "The genome sequence of the entomopathogenic bacterium Photobacterium
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; BX571868; CAE15179.1; -; Genomic_DNA.
DR Photobact; plu2805; -.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 208 AA; 22831 MW; 3B484FAP1BFB36AD CRC64;
Query Match 84.0%; Score 89; DB 1; Length 208;
Best Local Similarity 85.7%; Pred. No. 2e-06;
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 2 YDITLALAGICQSARLVQOLA 22
Db 6 YNITLALAGICQSARLVQOLA 26
RESULT 15
Y638.HABIN
ID Y638.HABIN STANDARD; PRT; 205 AA.
AC P44796;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein HI0638.
GN OrderedLocustNames=HI0638;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirtness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom B., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georgagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; U32747; AAC22298.1; -; Genomic_DNA.
DR PIR; I64155; I64155.

DR TIGR; H10638; -.
DR HAMAP; MF_00665; -. 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 205 AA; 2325 MW; 81F31B016BFF2DA CRC64;

Query Match	77.4%	Score 82;	DB 1;	Length 205;
Best Local Similarity	76.2%	Pred. No. 3e-05;		
Matches 16;	Conservative 3;	Mismatches 2;	Indels 0;	Gaps 0;

```

QY      2 YDITLALAGICQ$ARLVQQLA 22
        :|||:||||:|||||
Db      5 HDIVLALAGVCQ$AKLVHQLA 25

```

Search completed: January 23, 2006, 10:40:57
Job time : 70.131 secs

THE FORD FOUNDATION

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:35:50 ; Search time 20.1747 seconds
(without alignments)
90.156 Million cell updates/sec

Title: US-10-501-838a-27
Perfect score: 106
Sequence: 1 MYDITLALAGVCSARLVQOLA 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_AA:*

- 1: /cgn2_6/prodata/1/1aa/5/COMB.pep:*
- 2: /cgn2_6/prodata/1/1aa/6/COMB.pep:*
- 3: /cgn2_6/prodata/1/1aa/H/COMB.pep:*
- 4: /cgn2_6/prodata/1/1aa/PCTUS/COMB.pep:*
- 5: /cgn2_6/prodata/1/1aa/RE/COMB.pep:*
- 6: /cgn2_6/prodata/1/1aa/backfillset.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	91.5	217	US-09-489-039A-7789	Sequence 7789, App
2	86	81.1	216	US-09-543-681A-7993	Sequence 7993, App
3	43	40.6	60	US-09-248-796A-27407	Sequence 27407, A
4	42.5	40.1	427	US-09-248-796A-17203	Sequence 17203, A
5	42.5	40.1	762	US-09-228-986-114	Sequence 114, App
6	42.5	39.6	833	US-09-252-991A-23655	Sequence 23655, A
7	42	38.7	114	US-09-902-540-10640	Sequence 10640, A
8	41	38.7	309	US-09-902-540-12828	Sequence 12828, A
9	41	38.7	311	US-09-489-039A-12640	Sequence 12640, A
10	41	38.7	499	US-09-902-540-14780	Sequence 14780, A
11	41	38.7	588	US-09-605-703B-280	Sequence 280, App
12	41	38.7	588	US-09-605-703B-282	Sequence 282, App
13	41	38.7	588	US-08-540-406-8	Sequence 8, App1
14	40	37.7	115	US-08-656-055-8	Sequence 8, App1
15	40	37.7	115	US-08-954-668-8	Sequence 8, App1
16	40	37.7	115	US-08-918-658-8	Sequence 8, App1
17	40	37.7	115	US-09-724-631-8	Sequence 8, App1
18	40	37.7	115	US-08-954-701A-8	Sequence 8, App1
19	40	37.7	115	US-09-754-032-8	Sequence 8, App1
20	40	37.7	115	US-08-916-140-8	Sequence 8, App1
21	40	37.7	115	PCT-US95-13233-8	Sequence 8, App1
22	40	37.7	115	PCT-US95-13233-8	Sequence 8, App1
23	40	37.7	138	US-09-270-767-58510	Sequence 58510, A
24	40	37.7	324	US-09-489-039A-10801	Sequence 10801, A
25	40	37.7	398	US-09-543-681A-7046	Sequence 7046, App
26	40	37.7	409	US-08-360-673-6	Sequence 6, App1
27	40	37.7	409	US-09-640-305-6	Sequence 6, App1

28	40	37.7	410	US-08-088-633-2	Sequence 2, App1
29	40	37.7	410	US-08-245-756-2	Sequence 2, App1
30	40	37.7	410	US-08-441-750-2	Sequence 2, App1
31	40	37.7	410	US-08-441-751-2	Sequence 2, App1
32	40	37.7	410	PCT-US92-02521-2	Sequence 2, App1
33	40	37.7	433	US-09-270-767-43172	Sequence 43172, A
34	40	37.7	440	US-09-248-796A-18383	Sequence 18383, A
35	40	37.7	453	US-09-252-991A-32168	Sequence 32168, A
36	40	37.7	576	US-09-248-796A-20352	Sequence 20352, A
37	40	37.7	1021	US-09-902-540-16773	Sequence 16773, A
38	39	36.8	94	US-09-198-452A-1181	Sequence 1181, App
39	39	36.8	231	US-09-830-433A-32	Sequence 32, App1
40	39	36.8	300	US-09-489-039A-13439	Sequence 13439, A
41	39	36.8	306	US-09-392-772-6	Sequence 6, App1
42	39	36.8	350	US-09-438-185A-285	Sequence 285, App
43	39	36.8	367	US-09-248-796A-19809	Sequence 19809, A
44	39	36.8	398	US-08-328-314-2	Sequence 2, App1
45	39	36.8	398	US-08-731-045-2	Sequence 2, App1

ALIGNMENTS

```
RESULT 1
US-09-489-039A-7789
; Sequence 7789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7789
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match          91.5%; Score 97; DB 2; Length 217;
Best Local Similarity 90.5%; Pred. No. 1.9e+08;
Matches 19; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 YDITLALAGVCSARLVQOLA 22
Db      10 YDITLALAGVCSARLVQOLA 30

RESULT 2
US-09-543-681A-7993
; Sequence 7993, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543, 681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7993
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7993

Query Match          81.1%; Score 86; DB 2; Length 216;
```

Best Local Similarity 85.0%; Pred. No. 13e-06;
Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 DITLALAGIOSARLVQOLA 22
| | | | | | | | | | | | | | | | | | | | | |
Db 14 DITLALAGIOSARLVQOLA 33

RESULT 3

US-09-248-796A-27407
; Sequence 27407, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 27407
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-27407

Query Match 40.6%; Score 43; DB 2; Length 60;
Best Local Similarity 50.0%; Pred. No. 5.4;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MYDITLALAGIOS 14
| | | | | | | | | | | | | | | | | | | | | |
Db 43 LYDINMIIIVGICLS 56

RESULT 4

US-09-248-796A-17203
; Sequence 17203, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17203
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (339)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-17203

Query Match 40.1%; Score 42.5; DB 2; Length 427;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 14; Conservative 1; Mismatches 6; Indels 7; Gaps 2;

Qy 2 YDITLALAG--IC-----OSARLVQOLA 22
| | | | | | | | | | | | | | | | | | | | | |
Db 120 YDILKTIQVPGICGICGQSASLVQOLA 147

RESULT 5
US-09-228-986-114
; Sequence 114, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Nieuwenhuizen, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 114
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-114

Query Match 40.1%; Score 42.5; DB 2; Length 762;
Best Local Similarity 70.6%; Pred. No. 1.1e+02;
Matches 12; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 7 ALAGIOSARLVQOLA 22
| | | | | | | | | | | | | | | | | | | | | |
Db 387 ALSSILQASRLNVEOLA 403

RESULT 6

US-10-101-464A-114
; Sequence 114, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-114

Query Match 40.1%; Score 42.5; DB 2; Length 762;
Best Local Similarity 70.6%; Pred. No. 1.1e+02;
Matches 12; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 7 ALAGIOSARLVQOLA 22
| | | | | | | | | | | | | | | | | | | | | |
Db 387 ALSSILQASRLNVEOLA 403

RESULT 7
US-09-252-991A-23655
; Sequence 23655, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:


```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23655
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23655

Query Match      39.6%; Score 42; DB 2; Length 823;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      8 LAGICOSARLVQOLA 22
DB      524 LAGICQAPRLQAQAA 538

RESULT 8
US-09-902-540-10640
; Sequence 10640, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10640
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10640

Query Match      38.7%; Score 41; DB 2; Length 114;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      6 LALAGICOSARLVQ 19
DB      12 VVLGACNVARLIQ 25

RESULT 9
US-09-902-540-12828
; Sequence 12828, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825

```

```

; SEQ ID NO 12828
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12828

Query Match      38.7%; Score 41; DB 2; Length 309;
Best Local Similarity 56.2%; Pred. No. 74;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      7 ALAGICOSARLVQOLA 22
DB      152 ALAGVQERARLFHSFA 167

RESULT 10
US-09-489-039A-12640
; Sequence 12640, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12640
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12640

Query Match      38.7%; Score 41; DB 2; Length 311;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      1 MYDITLALAGICOSAR 16
DB      86 VYDLTKVMSIRQAAAR 101

RESULT 11
US-09-902-540-14780
; Sequence 14780, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14780
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14780

Query Match      38.7%; Score 41; DB 2; Length 499;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      8 LAGICOSARLVQOLA 22
DB      377 LAGFVQARLRPRRA 391

```

RESULT 12
US-09-605-703B-280
; Sequence 280, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompeju, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 280
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-280

Query Match 38.7%; Score 41; DB 2; Length 588;
Best Local Similarity 42.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 YDITLALGICOSARLVQ 20
DB 521 FDIAGGAGISEDAVLTQ 539

RESULT 13
US-09-605-703B-282
; Sequence 282, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompeju, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 282
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-282

Query Match 38.7%; Score 41; DB 2; Length 588;
Best Local Similarity 42.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 YDITLALGICOSARLVQ 20
DB 521 FDIAGGAGISEDAVLTQ 539

RESULT 14

US-08-540-406-8
; Sequence 8, Application US/08540406
; Patent No. 5837538
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Berttram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-761-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 115 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-540-406-8

Query Match 37.7%; Score 40; DB 1; Length 115;
Best Local Similarity 57.1%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MYDITLALGICOS 14
DB 69 MYDITWLKIDMCYS 82

RESULT 15
US-08-656-055-8
; Sequence 8, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/656,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/540,406
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 115 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: peptide
US-08-656-055-8

Query Match 37.7%; Score 40; DB 2; Length 115;
Best Local Similarity 57.1%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MYDITLAGICQS 14
DB 69 MYDITWIKDKMCT 82

Search completed: January 23, 2006, 10:46:50
Job time: 20.1747 secs

THE GOOD LIFE
A JOURNAL OF
LIVING

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using SW model

Run on: January 23, 2006, 11:53:39 ; Search time 3.07424 Seconds
(without alignments)
72.521 Million cell updates/sec

Title: US-10-501-838a-27

Sequence: 1 MYDITLALGICQSARLVQOLA 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1: Published Applications AA, New: *
2: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppa/US05_NEW_PUB.pep:*
7: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	42.5	185	6	US-10-453-372-896
2	45	42.5	194	6	US-10-453-372-894
3	45	42.5	218	6	US-10-453-372-892
4	43	40.6	42	6	US-10-512-295A-3
5	41.5	39.2	365	7	US-11-000-463-241
6	41.5	39.2	365	7	US-11-000-463-713
7	41	38.7	305	6	US-10-520-820-1
8	40	37.7	203	6	US-10-453-372-890
9	40	37.7	1015	6	US-10-957-569-9
10	38	35.8	211	6	US-10-965-972-4
11	38	35.8	211	7	US-11-080-991-84
12	38	35.8	655	6	US-10-055-877-40
13	38	35.8	678	6	US-10-055-877-202
14	38	35.8	687	6	US-10-055-877-199
15	38	35.8	687	6	US-10-055-877-201
16	38	35.8	919	6	US-10-858-730-206
17	37	34.9	367	6	US-10-821-234-1569
18	36	34.0	228	6	US-10-467-557-1242
19	36	34.0	250	6	US-10-793-626-2310
20	36	34.0	301	6	US-10-467-557-7356
21	36	34.0	434	6	US-10-467-557-6894
22	35.5	33.5	341	6	US-10-467-557-154
23	35.5	33.5	341	6	US-10-467-557-5822
24	35.5	33.5	341	6	US-10-467-557-7476
25	35.5	33.5	1259	6	US-10-467-557-5510

26	35	33.0	293	6	US-10-967-671-15	Sequence 15, Appl
27	35	33.0	332	6	US-10-454-437-108	Sequence 308, App
28	35	33.0	356	6	US-10-793-626-2054	Sequence 2054, Ap
29	35	33.0	445	7	US-11-194-246-345	Sequence 345, App
30	35	33.0	502	6	US-10-063-703-158	Sequence 158, App
31	35	33.0	502	7	US-11-102-240-158	Sequence 158, App
32	35	33.0	651	7	US-11-124-368A-184	Sequence 184, Appl
33	34.5	32.5	1128	7	US-11-037-243-97	Sequence 97, Appl
34	34	32.1	138	7	US-11-051-453-56	Sequence 56, Appl
35	34	32.1	161	6	US-10-821-234-1124	Sequence 1124, Ap
36	34	32.1	433	6	US-10-131-826A-334	Sequence 334, App
37	34	32.1	474	7	US-11-000-463-284	Sequence 284, App
38	34	32.1	687	6	US-10-055-877-200	Sequence 200, App
39	34	32.1	953	7	US-11-037-243-66	Sequence 66, Appl
40	34	32.1	1236	7	US-11-080-991-68	Sequence 68, Appl
41	34	32.1	2715	7	US-11-113-424-51	Sequence 51, Appl
42	33.5	31.6	403	6	US-10-763-712A-27	Sequence 27, Appl
43	33.5	31.6	449	6	US-10-763-712A-83	Sequence 83, Appl
44	33.5	31.6	468	6	US-10-467-557-776	Sequence 776, App
45	33.5	31.6	481	6	US-10-467-557-4660	Sequence 4660, Ap

ALIGNMENTS

RESULT 1
US-10-453-372-896
; Sequence 896, Application US/10453372
; Publication No. US20060003323A1
GENERAL INFORMATION:
; APPLICANT: Alsbjork, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT FILING DATE: 2003-06-03
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CurSeqRel version 0.1
; SEQ ID NO 896
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-896
Query Match 42.5%; Score 45; DB 6; Length 185;
Best Local Similarity 50.0%; Pred. No. 0.92;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MYDITLALGICQSARLVQOLA 22
: ||| ||| ||| :
Db 66 LYDSLALDGHIGSARLVQVVA 87

```
RESULT 2
US-10-453-372-894
; Sequence 894, Application US/10453372
; Publication No. US20060003323a1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Cnaseqdist version 0.1
; SEQ ID NO 894
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-894

Query Match      42.5%; Score 45; DB 6; Length 194;
Best Local Similarity 50.0%; Pred. No. 0.96;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Cy      1 MYDITLALGICQSARLVQOLA 22
Db      46 LYDSLALDGHIGSARALMVVA 67

RESULT 3
US-10-453-372-892
; Sequence 892, Application US/10453372
; Publication No. US20060003323a1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
```

```
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Cnaseqdist version 0.1
; SEQ ID NO 892
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-892

Query Match      42.5%; Score 45; DB 6; Length 218;
Best Local Similarity 50.0%; Pred. No. 1.1;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Cy      1 MYDITLALGICQSARLVQOLA 22
Db      70 LYDSLALDGHIGSARALMVVA 91

RESULT 4
US-10-512-295A-3
; Sequence 3, Application US/10512295A
; Publication No. US20050245727a1
; GENERAL INFORMATION:
; APPLICANT: Balizer, Gunnar
; APPLICANT: Dolphin, Gunnar
; APPLICANT: Liedberg, Bo
; APPLICANT: Lundstrom, Ingemar
; TITLE OF INVENTION: NOVEL POLYPEPTIDE SCARFOLDS AND USE THEREOF
; FILE REFERENCE: 5848.181USWO
; CURRENT APPLICATION NUMBER: US/10/512,295A
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/SE03/00507
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: SE 0200968-6
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide XE3
US-10-512-295A-3

Query Match      40.6%; Score 43; DB 6; Length 42;
Best Local Similarity 52.9%; Pred. No. 0.41;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy      6 LALAGICQSARLVQOLA 22
Db      16 LAAGPCDAQAIAEQOLA 32

RESULT 5
US-11-000-463-241
; Sequence 241, Application US/11000463
; Publication No. US20050266423a1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhilwei
; APPLICANT: Weinman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
```

```
/ APPLICANT: Cao, Yi-Cheng
/ APPLICANT: Dmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
/ FILE REFERENCE: 785CIP4CN
/ CURRENT APPLICATION NUMBER: US/11/000,463
/ PRIOR FILING DATE: 2004-11-29
/ PRIOR APPLICATION NUMBER: 10/291,265
/ PRIOR FILING DATE: 2002-11-08
/ PRIOR APPLICATION NUMBER: PCT/US01/02623
/ PRIOR FILING DATE: 2001-01-25
/ PRIOR APPLICATION NUMBER: 09/922,279
/ PRIOR FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: 09/491,404
/ PRIOR FILING DATE: 2000-01-25
/ PRIOR APPLICATION NUMBER: 09/617,746
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 09/631,451
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 09/633,870
/ NUMBER OF SEQ ID NOS: 944
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 241
/ LENGTH: 365
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-000-463-241

Query Match          39.2%; Score 41.5; DB 7; Length 365;
Best Local Similarity 40.0%; Pred. No. 7.5;
Matches 10; Conservative 4; Mismatches 8; Indels 3; Gaps 1;

QY      1 MYDITLALAGIC---QSARLVQOLA 22
         |||:|||||:|:|:|
Db      65 LIDMIGLAGFCGSMASGHILFKQMA 89

RESULT 6
US-11-000-463-713
/ Sequence 713, Application US/11000463
/ Publication No. US2005026423A1
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Adundi, Vinod
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Qian, Xiaohong B.
/ APPLICANT: Wang, Zhilwei
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Zhang, Jie
/ APPLICANT: Zhou, Ping
/ APPLICANT: Cao, Yi-Cheng
/ APPLICANT: Dmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
/ FILE REFERENCE: 785CIP4CN
/ CURRENT APPLICATION NUMBER: US/11/000,463
/ PRIOR FILING DATE: 2004-11-29
/ PRIOR APPLICATION NUMBER: 10/291,265
/ PRIOR FILING DATE: 2002-11-08
/ PRIOR APPLICATION NUMBER: PCT/US01/02623
/ PRIOR FILING DATE: 2001-01-25
/ PRIOR APPLICATION NUMBER: 09/922,279
/ PRIOR FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: 09/491,404
/ PRIOR FILING DATE: 2000-01-25
/ PRIOR APPLICATION NUMBER: 09/617,746
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 09/631,451
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 09/633,870
/ NUMBER OF SEQ ID NOS: 944
/ SOFTWARE: FastSeq for Windows Version 3.0
```

```
/ SEQ ID NO 713
/ LENGTH: 365
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-000-463-713

Query Match          39.2%; Score 41.5; DB 7; Length 365;
Best Local Similarity 40.0%; Pred. No. 7.5;
Matches 10; Conservative 4; Mismatches 8; Indels 3; Gaps 1;

QY      1 MYDITLALAGIC---QSARLVQOLA 22
         |||:|||||:|:|:|
Db      65 LIDMIGLAGFCGSMASGHILFKQMA 89

RESULT 7
US-10-520-820-1
/ Sequence 1, Application US/10520820
/ Publication No. US20060003393A1
/ GENERAL INFORMATION:
/ APPLICANT: MUTABILIS S.A.
/ TITLE OF INVENTION: Pathogenicity determinants which can be used as targets for devel
/ TITLE OF INVENTION: means for preventing and controlling bacterial infections and/or
/ TITLE OF INVENTION: dissemination
/ FILE REFERENCE: 1621
/ CURRENT APPLICATION NUMBER: US/10/520,820
/ CURRENT FILING DATE: 2005-01-07
/ PRIOR APPLICATION NUMBER: PCT/EP2003/008209
/ PRIOR FILING DATE: 2003-07-09
/ PRIOR APPLICATION NUMBER: FR 0208636
/ PRIOR FILING DATE: 2002-07-09
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 305
/ TYPE: PRT
/ ORGANISM: Escherichia coli
US-10-520-820-1
```

```
Query Match          38.7%; Score 41; DB 6; Length 305;
Best Local Similarity 42.9%; Pred. No. 7.5;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY      2 YDITLALAGICQSARLVQOLA 22
         |||:|||||:|:~|:|
Db      69 YDAVIDAGILVKSALVTRLA 89

RESULT 8
US-10-453-372-890
/ Sequence 890, Application US/10453372
/ Publication No. US20060003323A1
/ GENERAL INFORMATION:
/ APPLICANT: Alsebrook, et al.
/ TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHO
/ FILE REFERENCE: 21402-589 A
/ CURRENT APPLICATION NUMBER: US/10/453,372
/ CURRENT FILING DATE: 2003-06-03
/ PRIOR APPLICATION NUMBER: 09/789390
/ PRIOR FILING DATE: 2001-02-23
/ PRIOR APPLICATION NUMBER: 60/185967
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 09/823187
/ PRIOR FILING DATE: 2001-03-29
/ PRIOR APPLICATION NUMBER: 60/195792
/ PRIOR FILING DATE: 2000-03-10
/ PRIOR APPLICATION NUMBER: 09/839446
/ PRIOR FILING DATE: 2001-03-19
/ PRIOR APPLICATION NUMBER: 60/199476
/ PRIOR FILING DATE: 2000-03-25
/ PRIOR APPLICATION NUMBER: 09/863776
/ PRIOR FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: 60/208263
```

```

; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Cnaseqdist version 0.1
; SEQ ID NO: 890
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-890

```

```

Query Match      37.7%; Score 40; DB 6; Length 203;
Best Local Similarity 45.5%; Pred. No. 7.2;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

```

```

Qy      1 MYDITLALAGICQSARLVQOLA 22
Db      66 LYDSLALDGRQARALMVVA 87

```

```

RESULT 9
US-10-957-569-51
; Sequence 51, Application US/10957569
; Publication No. US20050246785A1
; GENERAL INFORMATION:

```

```

; APPLICANT: COOK, Zhihong et al.
; TITLE OF INVENTION: PROMOTER, PROMOTER CONTROL ELEMENTS, AND COMBINATIONS, AND USES
; FILE REFERENCE: 2750-1577PUS3
; CURRENT APPLICATION NUMBER: US/10/957,569
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 10/950,321
; PRIOR FILING DATE: 2004-09-23
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 51
; LENGTH: 1015
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-957-569-51

```

```

Query Match      37.7%; Score 40; DB 6; Length 1015;
Best Local Similarity 31.8%; Pred. No. 40;
Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy      1 MYDITLALAGICQSARLVQOLA 22
Db      170 LQDWTLMITGVCAPFSLIVGIA 191

```

```

RESULT 10
US-10-965-972-4
; Sequence 4, Application US/10965972
; Publication No. US20050266421A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bitd, Timothy A.
; APPLICANT: Youakim, Adel
; TITLE OF INVENTION: Claudin Polypeptides, Polynucleotides, and Methods of Making and
; FILE REFERENCE: 3426-NO
; CURRENT APPLICATION NUMBER: US/10/965,972
; CURRENT FILING DATE: 2004-10-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 4
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-965-972-4

```

```

Query Match      35.8%; Score 38; DB 6; Length 211;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy      1 MYDITLALAGICQSAR 16
Db      66 MYDSVLAALSAALQATR 81

```

```

RESULT 11
US-11-080-991-84
; Sequence 84, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Velby, Pelter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 84
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-080-991-84

```

```

Query Match      35.8%; Score 38; DB 7; Length 211;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy      1 MYDITLALAGICQSAR 16
Db      66 MYDSVLAALSAALQATR 81

```

```

RESULT 12
US-10-055-877-40
; Sequence 40, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kerkuda, Rameesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zernuusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Paturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shimkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li

```


APPLICANT: Casman, Stacie
APPLICANT: Bollog, Ference
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 655
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-877-40

Query Match 35.8%; Score 38; DB 6; Length 655;
Best Local Similarity 64.3%; Pred. No. 55;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 LAGICQARLVQOL 21
DB 558 LVGIVQRAQLVQAL 571

RESULT 13
US-10-055-877-202
Sequence 202, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zerhusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patuturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eissen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shimkets, Richard
APPLICANT: Gusev, Vladimr
APPLICANT: Verne, Corine
APPLICANT: Taupier Jr., Raymond

APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Bollog, Ference
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 202
LENGTH: 678
TYPE: PRT
ORGANISM: Oryzctolagus cuniculus
US-10-055-877-202

Query Match 35.8%; Score 38; DB 6; Length 678;
Best Local Similarity 64.3%; Pred. No. 57;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 LAGICQARLVQOL 21
DB 590 LVGIVQRAQLVQAL 603

RESULT 14
US-10-055-877-199
Sequence 199, Application US/10055877
Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zerhusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patuturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eissen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shimkets, Richard

```

; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Caeman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; PRIORITY FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 199
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-877-199

Query Match          35.8%; Score 38; DB 6; Length 687;
Best Local Similarity 64.3%; Pred. No. 58;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy      8 LAGICQSARLVQQL 21
Db      590 LVGIQRAQLVQAL 603

RESULT 15
US-10-055-877-201
; Sequence 201, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Keskula, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zehnusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patuturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Biesen, Andrew
```

```

; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shinkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Caeman, Stacie
; APPLICANT: Boldog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; PRIORITY FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 201
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Oryzctolagus cuniculus
US-10-055-877-201

Query Match          35.8%; Score 38; DB 6; Length 687;
Best Local Similarity 64.3%; Pred. No. 58;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy      8 LAGICQSARLVQQL 21
Db      590 LVGIQRAQLVQAL 603
```

Search completed: January 23, 2006, 12:09:43
Job time : 4.07424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:20:59 ; Search time 73.0131 seconds
(without alignments)
132.392 Million cell updates/sec

Title: US-10-501-838a-27
Sequence: 1 MYDITLALGICQSARLVQQLA 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Genesegp19808:*
2: Genesegp19908:*
3: Genesegp20008:*
4: Genesegp20018:*
5: Genesegp20028:*
6: Genesegp20038:*
7: Genesegp20048:*
8: Genesegp20058:*
9: Genesegp20068:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	22	ADB16910	Adb16910 Escherich
2	106	100.0	22	AEBO8244	Aeb08244 Escherich
3	106	100.0	29	ADB16920	Adb16920 Escherich
4	106	100.0	29	AEBO8254	Aeb08254 Penetrati
5	101	95.3	23	ADB16888	Adb16888 Escherich
6	101	95.3	23	ADB16909	Adb16909 Escherich
7	101	95.3	23	AEBO8220	Aeb08220 Escherich
8	101	95.3	23	AEBO8243	Aeb08243 Escherich
9	101	95.3	24	ADB16908	Adb16908 Escherich
10	101	95.3	24	AEBO8242	Aeb08242 Escherich
11	101	95.3	25	ADB16922	Adb16922 Escherich
12	101	95.3	25	ADB16906	Adb16906 Escherich
13	101	95.3	26	ADB16904	Adb16904 E.coli pe
14	101	95.3	29	ADB16921	Adb16921 E.coli pe
15	101	95.3	30	ADB16903	Adb16903 E.coli pe
16	101	95.3	30	ADB16919	Adb16919 Escherich
17	101	95.3	30	ADB16905	Adb16905 Escherich
18	101	95.3	30	AEBO8239	Aeb08239 Penetrati
19	101	95.3	30	AEBO8253	Aeb08253 Penetrati
20	101	95.3	31	ADB16918	Adb16918 Escherich
21	101	95.3	31	AEBO8252	Aeb08252 Penetrati
22	101	95.3	213	AEBO8263	Aeb08263 Escherich
23	97	91.5	217	ABO61272	Ab061272 Klebsiell
24	90	84.9	23	ADB16887	Adb16887 Pasteurel

25	90	84.9	23	9	AEBO8219	Aeb08219 Pasteurel
26	90	84.9	203	6	AEBO8262	Aeb08262 Pasteurel
27	89	84.0	209	6	ABW67069	Abw67069 Photorhab
28	86	81.1	216	7	ADOF0708	Adof0708 Bacteri
29	85	80.2	23	6	ADB16912	Adb16912 Haemophil
30	85	80.2	23	6	AEBO8246	Aeb08246 Haemophil
31	85	80.2	29	6	ADB16916	Adb16916 Escherich
32	85	80.2	29	6	AEBO8250	Aeb08250 Penetrati
33	82	77.4	23	6	ADB16886	Adb16886 Haemophil
34	82	77.4	23	6	AEBO8218	Aeb08218 Haemophil
35	82	77.4	205	4	AAW50230	Aaw50230 Haemophil
36	82	77.4	205	4	AEBO8261	Aeb08261 Buchnera
37	80	75.5	23	6	ADB16890	Adb16890 Buchnera
38	80	75.5	23	9	AEBO8222	Aeb08222 Buchnera
39	80	75.5	23	9	AEBO8265	Aeb08265 Vibrio ch
40	73	68.9	23	6	ADB16889	Adb16889 Vibrio ch
41	73	68.9	23	9	AEBO8221	Aeb08221 Vibrio ch
42	73	68.9	204	9	AEBO8264	Aeb08264 Rhizobium
43	48	45.3	23	6	ADB16896	Adb16896 Mesorhizo
44	48	45.3	23	9	AEBO8228	Aeb08228 Mesorhizo
45	46	43.4	223	6	ABU21532	Abu21532 Protein e

ALIGNMENTS

RESULT 1
ADB16910
ID ADB16910 standard; peptide; 22 AA.

XX ADB16910;

XX 20-NOV-2003 (first entry)

XX Escherichia coli YCFC penetrating peptide 32.

XX Penetrating peptide; epithelial; endothelial; tight junction; diabetes;
XX infertility; hormone; vitamin deficiency; neurodegenerative;
XX cardiovascular; haematological; endocrine disorder; obesity;
XX neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
XX osteopathic; cytostatic; nootropic.

XX Escherichia coli.

XX WO2003066859-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003WO-1B000968.

XX 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Ben-Sasson SA, Cohen B;

XX WPI; 2003-697452/66.

XX New penetrating peptide, useful for preparing a composition for treating

XX or preventing e.g. endocrine disorders.

XX Claim 2; Page 15; 60pp; English.

XX This invention relates to a novel peptide sequences capable of

XX translocating across a biological barrier. Furthermore, it refers to

XX methods that use these peptides to facilitate penetration of a

XX biologically active effector molecule such as a drug or other therapeutic

XX agent across biological barriers e.g. epithelial or endothelial cells

XX sealed by tight junctions. This peptide is derived from a bacterial

XX toxin, an integral membrane or extracellular protein and can comprise an

XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin

XX or enzyme. The effector molecule, however, can comprise for example

XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or nocotropic activities. This peptide is from YCFC of
 CC Becherichia coli and is penetrating peptide 32 of the invention.
 XX
 SQ Sequence 22 AA;
 Query Match 100.0%; Score 106; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 2.7e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MYDITLALAGICQSARLVQQLA 22
 Db 1 MYDITLALAGICQSARLVQQLA 22
 RESULT 2
 AEB08244
 ID AEB08244 standard; peptide; 22 AA.
 AC AEB08244;
 XX
 DT 25-AUG-2005 (first entry)
 DE Becherichia coli YCFC penetrating peptide 32, SEQ ID NO: 27.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease; diabetes;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW nocotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiatherosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antihaemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antipneumatic; cyostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Becherichia coli.
 XX
 PN US2005136103-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 16-SEP-2004; 2004US-00942300.
 XX
 PR 17-SEP-2003; 2003US-00664989.
 PR 17-SEP-2003; 2003US-00665184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 XX
 PI Ben-Sasson SA, Cohen E;
 XX
 DR WPI, 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 27; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Becherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 SQ Sequence 22 AA;
 Query Match 100.0%; Score 106; DB 9; Length 22;
 Best Local Similarity 100.0%; Pred. No. 2.7e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MYDITLALAGICQSARLVQQLA 22
 Db 1 MYDITLALAGICQSARLVQQLA 22
 RESULT 3
 ADB16920
 ID ADB16920 standard; peptide; 29 AA.
 AC ADB16920;
 XX
 DT 20-NOV-2003 (first entry)
 DE Becherichia coli peptide 3 coupled to imaging compound linker, IBW-002V3.
 XX
 KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiac; antiatherosclerotic; osteopathic; cyostatic; nocotropic;
 KW imaging linker; penetrating peptide; IBW-002V3.
 XX
 OS Synthetic.
 OS Becherichia coli.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 29 /note= "N-terminal acetyl"
 FT Modified-site 29 /note= "C-terminal amide"
 XX
 PN WO2003066859-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 07-FEB-2003; 2003WO-IB000968.
 XX
 PR 07-FEB-2002; 2002US-0355396P.
 XX

PA (Y188) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 PI
 XX WPI; 2003-697452/66.
 DR
 XX
 PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 PS Example 3; Page 40; 60pp; English.
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, daltargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiant, antiarteriosclerotic, osteoporotic,
 CC cytoskeletal or neurotrophic activities. This peptide sequence is IBM-002V3,
 CC which consists of the Escherichia coli penetrating peptide 3 coupled to
 CC the imaging linker peptide used in an exemplification of the invention.
 XX
 SQ Sequence 29 AA;
 Query Match 100.0%; Score 106; DB 6; Length 29;
 Best Local Similarity 100.0%; Pred. No. 3.7e-10; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 0;
 QY 1 MYDITLALAGICQSARLVQOLA 22
 DB 1 MYDITLALAGICQSARLVQOLA 22
 RESULT 4
 AEB08254 standard; peptide; 29 AA.
 AC AEB08254;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DB Penetrating peptide SEQ: 37 used in composition for mucosal vaccination.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimers disease;
 KW neurotrophic; neurological disease; parkinsons disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiant; vasotrophic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephroretropic;
 KW genitourinary disease; hematological disease; antiemetic; anemia;
 KW auto-immune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytoskeletal;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.

XX Unidentified.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT Misc-difference 1
 FT Misc-difference 1
 FT Misc-difference 29
 FT Misc-difference 29
 FT amino group of lysine is acylated with a fatty acid"
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHE/) COHEN E.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 63; SEQ ID NO 37; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes;
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological,
 CC disorders, neurodegenerative disorders, Alzheimers disease,
 CC parkinsons disease, multiple sclerosis, Huntingtons disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is a penetrating peptide (IBM-002V3)
 CC used in the composition for mucosal vaccination using a counter ion and a
 CC penetrating peptide.
 XX
 SQ Sequence 29 AA;
 Query Match 100.0%; Score 106; DB 9; Length 29;
 Best Local Similarity 100.0%; Pred. No. 3.7e-10; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 0;
 QY 1 MYDITLALAGICQSARLVQOLA 22
 DB 1 MYDITLALAGICQSARLVQOLA 22
 RESULT 5
 ADB16888 standard; peptide; 23 AA.
 AC ADB16888;
 XX

XX	20-NOV-2003	(first entry)
DT		
DE	Escherichia coli YCFC penetrating peptide 3.	
KW	penetrating peptide; epithelial; endothelial; tight junction; diabetes;	
KM	infertility; hormone; vitamin deficiency; neurodegenerative;	
KW	cardiovascular; haematological; endocrine disorder; obesity;	
KM	neoplastic disease; neuroprotective; cardiant; antiatherosclerotic;	
KW	osteopathic; cyostatic; nootropic.	
OS	Escherichia coli.	
XX		
PN	WO200306859-A2.	
XX		
PD	14-AUG-2003.	
XX		
PF	07-FEB-2003; 2003WO-IB000968.	
XX		
PR	07-FEB-2002; 2002US-0355396P.	
XX		
PA	(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.	
XX		
PL	Ben-Sasson SA, Cohen E;	
XX		
DR	WPI; 2003-697452/66.	
XX		
PT	New penetrating peptide, useful for preparing a composition for treating	
PT	or preventing e.g. endocrine disorders.	
XX		
PS	Claim 2; Page 14; 60pp; English.	
XX		
CC	This invention relates to a novel peptide sequences capable of	
CC	translocating across a biological barrier. Furthermore, it refers to	
CC	methods that use these peptides to facilitate penetration of a	
CC	biologically active effector molecule such as a drug or other therapeutic	
CC	agent across biological barriers e.g. epithelial or endothelial cells	
CC	sealed by tight junctions. This peptide is derived from a bacterial	
CC	toxin, an integral membrane or extracellular protein and can comprise an	
CC	anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin	
CC	or enzyme. The effector molecule, however, can comprise for example	
CC	insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony	
CC	stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic	
CC	factors. The penetrating peptide is useful for the treatment of various	
CC	conditions including diabetes, infertility, hormone and vitamin	
CC	deficiencies, neurodegenerative, cardiovascular, haematological and	
CC	endocrine disorders, as well as obesity and neoplastic disease.	
CC	Accordingly, the peptides of this invention can be used in compositions	
CC	that have neuroprotective, cardiant, antiatherosclerotic, vitamin,	
CC	cyostatic or nootropic activities. This peptide is from YCFC of	
CC	Escherichia coli and is penetrating peptide 3 of the invention.	
XX		
SQ	Sequence 23 AA;	
XX		
Query Match	95.3%; Score 101; DB 6; Length 23;	
Best Local Similarity	100.0%; Pred. No. 1.9e-09;	
Matches 21; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
OY	2 YDITLALAGICQSAARLVQGLA 22	
DB	3 YDITLALAGICQSAARLVQGLA 23	
XX		
RESULT 6		
ADBI6909		
ID	ADBI6909 standard; peptide; 23 AA.	
XX		
AC	ADBI6909;	
XX		
DT	20-NOV-2003 (first entry)	
XX		
DE	Escherichia coli YCFC penetrating peptide 31.	
XX		

KM		penetrating peptide; epithelial; endothelial; tight junction; diabetes;
KW		fertility; hormone; vitamin deficiency; neurodegenerative;
KX		cardiovascular; haematological; endocrine disorder; obesity;
KV		neoplastic disease; neuroprotective; candidant; antiarteriosclerotic;
KW		osteopathic; cyostatic; nootropic.
XX		
OS	Bescherichia coli.	
PN	WO200306859-A2.	
PD	14-AUG-2003.	
XX		
PF	07-FEB-2003; 2003WO-IB000968.	
PR	07-FEB-2002; 2002US-0355396P.	
PA	(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.	
FJ	Ben-Sasson SA, Cohen E;	
DR	WPI; 2003-697452/66.	
PT	New penetrating peptide, useful for preparing a composition for treating or preventing e.g. endocrine disorders.	
PS	Claim 2; Page 15; 60pp; English.	
CC	This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/macocyte colony stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, hematological and endocrine disorders, as well as obesity and neoplastic disease. CC Accordingly, the peptides of this invention can be used in compositions that have neuroprotective, candidant, antiarteriosclerotic, osteopathic, cytostatic or nootropic activities. This peptide is from YCPC of Bescherichia coli and is penetrating peptide 31 of the invention.	
SQ	Sequence 23 AA:	
Query Match	95.3%; Score 101; DB 6; Length 23;	
Best Local Similarity	100.0%; Pred. No. 1.9e-09;	
Matches	21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
DG	2 YDITLALAGICQSARLVQQLA 22 	
ID	3 YDITLALAGICQSARLVQQLA 23 	
AC	AEB08220 standard; peptide; 23 AA.	
AE	AEB08220;	
DT	25-AUG-2005 (first entry)	
DE	Escherichia coli YCPC penetrating peptide 3, SEQ ID NO: 3.	
XK	Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes; antidiabetic; endocrine disease; gastrointestinal diseases; metabolic disorder; hormone deficiency; osteoporosis; osteopathic; degeneration; musculoskeletal disease; ocular disease; ophthalmological; neurodegenerative disease; neuroprotective; Alzheimers disease;	

Db 4 NYRITLALAGICOSAHILWOLA 26

Search completed: January 23, 2006, 11:31:52
Job time : 160 Secs

This Page Blank (uspio)

KW noctropic; neurological disease; parkinsons disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antihypertensive;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; viroside;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 OS Escherichia coli.
 XX US2005136103-A1.
 XX 23-JUN-2005.
 XX 16-SEP-2004; 2004US-00942300.
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 DR MPI; 2005-444089/45.
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX Claim 53; SEQ ID NO 3; 59pp; English.
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against YCPFC and
 CC hepatitis B. The present sequence is the Escherichia coli YCPFC
 CC of antigenlycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX Sequence 23 AA;
 S0 Query Match 95.3%; Score 101; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 YDITLALGICQSARLVQOLA 23
 RESULT 8
 AEB08243
 ID AEB08243 standard; peptide; 23 AA.
 XX AEB08243;
 XX 25-AUG-2005 (first entry)
 DE Escherichia coli YCPFC penetrating peptide 31, SEQ ID NO: 26.
 KW pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW noctropic; neurological disease; parkinsons disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antihypertensive;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; viroside;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 OS Escherichia coli.
 XX US2005136103-A1.
 XX 23-JUN-2005.
 XX 16-SEP-2004; 2004US-00942300.
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 PA (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 DR MPI; 2005-444089/45.
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX Claim 46; SEQ ID NO 26; 59pp; English.
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic

CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC hematological disorders, anemia of different entities, immunologic and
CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
CC infectious diseases, viral infections, bacterial infections, fungal
CC infections, parasitic infections, neoplastic diseases, multi-factorial
CC disorders, impotence, chronic pain, depression, different fibrosis states
CC and short stature) and for mucosal vaccination against anthrax and
CC hepatitis B. The present sequence is the Escherichia coli YCFC
CC penetrating peptide. This sequence is used in the effective translocation
CC of aminoglycoside antibiotics and antifungal agents across an epithelial
CC barrier.
CC
XX
SQ Sequence 23 AA;
Query Match 95.3%; Score 101; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 YDITLALGICQSARLVQOLA 22
Db 3 YDITLALGICQSARLVQOLA 23
RESULT 9
ADBI6908
ID ADBI6908 standard; peptide; 24 AA.
AC ADBI6908;
XX
DT 20-NOV-2003 (first entry)
XX
DE Escherichia coli YCFC penetrating peptide 30.
XX
KW penetrating peptide; epithelial; endothelial; tight junction; diabetes;
KW infertility; hormone; vitamin deficiency; neurodegenerative;
KW cardiovascular; haematological; endocrine disorder; obesity;
KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
KW osteopathic; cyostatic; nootropic.
XX
OS Escherichia coli.
XX
PN WO2003066859-A2.
XX
PD 14-AUG-2003.
XX
PP 07-FEB-2003; 2003WO-IB000968.
XX
PR 07-FEB-2002; 2002US-0355396P.
XX
PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI Ben-Sasson SA, Cohen E;
XX
DR WPI; 2003-697452/66.
XX
PT New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
PS Claim 2; Page 15; 60pp; English.
XX
CC This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin

CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
CC cyostatic or nootropic activities. This peptide is from YCFC of
CC Escherichia coli and is penetrating peptide 30 of the invention.
XX
SQ Sequence 24 AA;
Query Match 95.3%; Score 101; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 YDITLALGICQSARLVQOLA 22
Db 4 YDITLALGICQSARLVQOLA 24
RESULT 10
AEB08242
ID AEB08242 standard; peptide; 24 AA.
AC AEB08242;
XX
DT 25-AUG-2005 (first entry)
XX
DE Escherichia coli YCFC penetrating peptide 30, SEQ ID NO: 25.
XX
KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
KW antidiabetic; endocrine disease; gastrointestinal disease;
KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
KW nootropic; neurological disease; parkinson's disease; antiparkinsonian;
KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
KW anticonvulsant; genetic disorder; cardiovascular disease;
KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
KW nutritional disorder; vitamin deficiency; renal disease; nephropathic;
KW genitourinary disease; hematological disease; antianemic; anemia;
KW autoimmune disease; immunosuppressive; immune deficiency;
KW immunostimulant; infectious disease; antimicrobial; infection;
KW erectile dysfunction; andrology; major depressive disorder;
KW antidepressant; psychiatric disorder; pain; analgesic;
KW bacterial infection; antibacterial; viral infection; vinticide;
KW fungal infection; fungicide; parasitic infection; antiparasitic;
KW renal failure; infertility; antirheumatic; cyostatic;
KW antiinflammatory; hepatotropic; hepatitis B virus infection.
XX
OS Escherichia coli.
XX
PN US2005136103-A1.
XX
PD 23-JUN-2005.
XX
PP 16-SEP-2004; 2004US-00942300.
XX
PR 17-SEP-2003; 2003US-00664989.
XX
PR 17-SEP-2003; 2003US-00665184.
XX
PA (BENS/) BEN-SASSON S A.
XX (COHE/) COHEN E.
XX
PI Ben-Sasson SA, Cohen E;
XX
DR WPI; 2005-444089/45.
XX
PT Composition used for translocating effectors across barrier such as
PT epithelial cells during treatment of e.g. endocrine disorders comprising
PT effector sequentially coupled with counter ion and hydrophobic agent.
XX
PS Claim 46; SEQ ID NO 25; 59pp; English.

XX The present invention relates to a pharmaceutical composition of
 CC present peptides for transendothelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC hemorological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YCFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.

XX Sequence 24 AA;

Query Match 95.3%; Score 101; DB 9; Length 24;
 Best Local Similarity 100.0%; Pred. No. 2e-03; 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0;

QY 2 YDITTLAAGICOSARLVQOLA 22
 |||||
 Db 4 YDITTLAAGICOSARLVQOLA 24

RESULT 11

ADBI6922
 ID ADBI6922 standard; peptide; 25 AA.

AC ADBI6922;

DT 20-NOV-2003 (first entry)

XX Escherichia coli peptide 3 linked to recombinant human insulin.

XX recombinant human insulin; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cytostatic; nootropic; penetrating peptide; diabetes.

XX Synthetic.
 OS Escherichia coli.

FT Key Location/Qualifiers
 FT Peptide 1..23
 FT Modified-site 25 /note= "Penetrating peptide 3"
 FT /note= "Recombinant human insulin peptide coupled to the
 C-terminus"

XX MO2003066859-A2.

XX 14-AUG-2003.

XX 07-FEB-2003; 2003MO-IB000968.

XX 07-FEB-2002; 2002US-0355396P.

XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

PI Ben-Sasson SA, Cohen E;

XX WPI; 2003-697452/66.

XX New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.

XX Example 8, Page 42; 60pp; English.

XX This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cytostatic or nootropic activities. This peptide sequence consists of the
 CC Escherichia coli penetrating peptide 3 coupled to recombinant human
 CC insulin in the absence of a detachable linker peptide as a cleavage site,
 CC used to deliver insulin across mucosal epithelia as a treatment for
 CC diabetes, in an exemplification of the invention.

XX Sequence 25 AA;

Query Match 95.3%; Score 101; DB 6; Length 25;
 Best Local Similarity 100.0%; Pred. No. 2.1e-09; 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0;

QY 2 YDITTLAAGICOSARLVQOLA 22
 |||||
 Db 3 YDITTLAAGICOSARLVQOLA 23

RESULT 12

ADBI6906
 ID ADBI6906 standard; peptide; 25 AA.

AC ADBI6906;

DT 20-NOV-2003 (first entry)

XX Escherichia coli peptide 3 linked to linearised insulin receptor peptide.

XX linearised insulin receptor; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cytostatic; nootropic; penetrating peptide; diabetes.

XX Synthetic.
 OS Escherichia coli.

FT Key Location/Qualifiers
 FT Peptide 1..23
 FT Modified-site 25 /note= "Penetrating peptide 3"
 FT /note= "linearised insulin receptor peptide coupled to
 the C-terminus"

XX MO2003066859-A2.

XX 14-AUG-2003.

PF 07-FEB-2003; 2003MO-IB000968.
XX
PR 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
PA
XX Ben-Sasson SA, Cohen E;
PI WPI; 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
XX Example 9; Page 43; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antihypertensive, osteopathic,
CC cyostatic or neurotropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to a linearised insulin
CC receptor, used to deliver insulin across mucosal epithelia as a treatment
CC for diabetes, in an exemplification of the invention.
XX
SQ Sequence 25 AA;

Query Match 95.3%; Score 101; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2 YDITLALAGICQSARLVQQLA 22
DB 3 YDITLALAGICQSARLVQQLA 23

RESULT 13
ADBI6904
ID ADBI6904 standard; peptide; 26 AA.
XX
AC ADBI6904;
XX
DT 20-NOV-2003 (first entry)
XX
DE E_coli penetrating peptide 3 linked to heparin without a cleavage site.
XX
XX heparin; epithelial; endothelial; tight junction; diabetes; infertility;
XX hormone; vitamin deficiency; neurodegenerative; cardiovascular;
XX haematological; endocrine disorder; obesity; neoplastic disease;
XX neuroprotective; cardiac; antihypertensive; osteopathic; cytostatic;
XX neurotropic; penetrating peptide.
XX
OS Synthetic.
XX Escherichia coli.
XX
XX Key Location/Qualifiers
XX Peptide 1..23
XX /note= "Penetrating peptide 3"
XX Modified-site 26
XX /note= "Heparin peptide coupled to the C-terminus"

PN W02003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003MO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
PA
XX Ben-Sasson SA, Cohen E;
PI WPI; 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
PT or preventing e.g. endocrine disorders.
XX
XX Example 10; Page 44; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiac, antihypertensive, osteopathic,
CC cyostatic or neurotropic activities. This peptide sequence consists of the
CC Escherichia coli penetrating peptide 3 coupled to heparin in the absence
CC of a detachable linker peptide as a cleavage site, used to deliver low
CC molecular weight heparin across mucosal epithelia in an exemplification
CC of the invention.
XX
SQ Sequence 26 AA;

Query Match 95.3%; Score 101; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2 YDITLALAGICQSARLVQQLA 22
DB 3 YDITLALAGICQSARLVQQLA 23

RESULT 14
ADBI6921
ID ADBI6921 standard; peptide; 29 AA.
XX
AC ADBI6921;
XX
DT 20-NOV-2003 (first entry)
XX
DE E_coli peptide 3 linked to recombinant human insulin via cleavage site.
XX
XX recombinant human insulin; epithelial; endothelial; tight junction;
XX diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
XX cardiovascular; haematological; endocrine disorder; obesity;
XX neoplastic disease; neuroprotective; cardiac; antihypertensive;
XX osteopathic; cytostatic; neurotropic; detachable linker;
XX penetrating peptide; diabetes.
XX
OS Synthetic.
XX Escherichia coli.

FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Cleavage-site /note= "Penetrating peptide 3"
 FT 26..29
 FT /note= "Detachable linker peptide"
 FT Modified-site 23
 FT /note= "Recombinant human insulin peptide coupled to the C-terminus"
 PN WO200306859-A2.
 PD 14-AUG-2003.
 PF 07-FEB-2003; 2003WO-IB000968.
 PR 07-FEB-2002; 2002US-0355396P.
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PI Ben-Sasson SA, Cohen E;
 DR WPI; 2003-697452/66.
 PT New penetrating peptide, useful for preparing a composition for treating or preventing e.g. endocrine disorders.
 PS Example 8; Page 42; 60pp; English.
 XX This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, haematological and endocrine disorders, as well as obesity and neoplastic disease. Accordingly, the peptides of this invention can be used in compositions that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic, cytoskeletal or neurotrophic activities. This peptide sequence consists of the Escherichia coli penetrating peptide 3 coupled to recombinant human insulin via the detachable linker peptide 3 as a cleavage site, used to deliver insulin across mucosal epithelia as a treatment for diabetes, in an exemplification of the invention.
 SQ Sequence 29 AA;
 Query Match 95.3%; Score 101; DB 6; Length 29;
 Best Local Similarity 100.0%; Pred. No. 2.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YDITLALAGICQSARLVQQLA 22
 DB 3 YDITLALAGICQSARLVQQLA 23
 RESULT 15
 ADB16903 standard; peptide; 30 AA.
 AC ADB16903;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE E.coli penetrating peptide 3 linked to heparin through a cleavage site.
 XX heparin; epithelial; endothelial; tight junction; diabetes; infertility;

KW hormone; vitamin deficiency; neurodegenerative; cardiovascular; haematological; endocrine disorder; obesity; neoplastic disease; KW neuroprotective; cardiac; antiarteriosclerotic; osteopathic; cytoskeletal; KW nototropic; detachable linker; penetrating peptide.
 OS Synthetic.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Cleavage-site /note= "Penetrating peptide 3"
 FT 26..29
 FT /note= "Detachable linker peptide"
 FT Modified-site 30
 FT /note= "Heparin peptide coupled to the C-terminus"
 PN WO200306859-A2.
 PD 14-AUG-2003.
 PF 07-FEB-2003; 2003WO-IB000968.
 PR 07-FEB-2002; 2002US-0355396P.
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PI Ben-Sasson SA, Cohen E;
 DR WPI; 2003-697452/66.
 PT New penetrating peptide, useful for preparing a composition for treating or preventing e.g. endocrine disorders.
 PS Example 10; Page 44; 60pp; English.
 XX This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, haematological and endocrine disorders, as well as obesity and neoplastic disease. Accordingly, the peptides of this invention can be used in compositions that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic, cytoskeletal or neurotrophic activities. This peptide sequence consists of the Escherichia coli penetrating peptide 3 coupled to heparin via the detachable linker peptide 3 as a cleavage site, used to deliver low molecular weight heparin across mucosal epithelia in an exemplification of the invention.
 SQ Sequence 30 AA;
 Query Match 95.3%; Score 101; DB 6; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YDITLALAGICQSARLVQQLA 22
 DB 3 YDITLALAGICQSARLVQQLA 23
 Search completed: January 23, 2006, 10:28:34
 Job time : 73.0131 secs

This logo mark (upside down)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 11:52:43 ; Search time 50.5327 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-27
Sequence: 1 MYDITLALAGICQSARLVQOLA 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA_Main:*

- 1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep:*
- 5: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*
- 6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	22	4	US-10-665-184-27 Sequence 27, Appl
2	106	100.0	22	5	US-10-942-300-27 Sequence 27, Appl
3	106	100.0	22	5	US-10-501-838a-27 Sequence 27, Appl
4	106	100.0	29	4	US-10-665-184-37 Sequence 37, Appl
5	106	100.0	29	5	US-10-942-300-37 Sequence 37, Appl
6	106	100.0	29	5	US-10-501-838a-37 Sequence 37, Appl
7	101	95.3	23	4	US-10-665-184-3 Sequence 3, Appl
8	101	95.3	23	4	US-10-665-184-26 Sequence 26, Appl
9	101	95.3	23	5	US-10-942-300-3 Sequence 3, Appl
10	101	95.3	23	5	US-10-942-300-26 Sequence 26, Appl
11	101	95.3	23	5	US-10-501-838a-3 Sequence 3, Appl
12	101	95.3	23	4	US-10-665-184-26 Sequence 26, Appl
13	101	95.3	24	5	US-10-942-300-25 Sequence 25, Appl
14	101	95.3	24	5	US-10-501-838a-25 Sequence 25, Appl
15	101	95.3	25	5	US-10-501-838a-19 Sequence 19, Appl
16	101	95.3	25	5	US-10-501-838a-23 Sequence 23, Appl
17	101	95.3	26	5	US-10-501-838a-21 Sequence 21, Appl
18	101	95.3	29	5	US-10-501-838a-18 Sequence 18, Appl
19	101	95.3	30	4	US-10-665-184-22 Sequence 22, Appl
20	101	95.3	30	4	US-10-665-184-36 Sequence 36, Appl
21	101	95.3	30	5	US-10-942-300-22 Sequence 22, Appl
22	101	95.3	30	5	US-10-942-300-36 Sequence 36, Appl
23	101	95.3	30	5	US-10-501-838a-20 Sequence 20, Appl
24	101	95.3	30	5	US-10-501-838a-22 Sequence 22, Appl
25	101	95.3	30	5	US-10-501-838a-36 Sequence 36, Appl
26	101	95.3	31	4	US-10-665-184-35 Sequence 35, Appl
27	101	95.3	31	4	US-10-665-184-35 Sequence 35, Appl

28	101	95.3	31	5	US-10-942-300-35 Sequence 35, Appl
29	101	95.3	31	5	US-10-501-838a-35 Sequence 35, Appl
30	101	95.3	213	5	US-10-665-184-61 Sequence 61, Appl
31	101	95.3	213	5	US-10-942-300-61 Sequence 61, Appl
32	101	95.3	213	5	US-10-501-838a-55 Sequence 55, Appl
33	90	84.9	23	4	US-10-665-184-2 Sequence 2, Appl
34	90	84.9	23	5	US-10-942-300-2 Sequence 2, Appl
35	90	84.9	23	5	US-10-501-838a-2 Sequence 2, Appl
36	90	84.9	203	5	US-10-665-184-60 Sequence 60, Appl
37	90	84.9	203	5	US-10-942-300-60 Sequence 60, Appl
38	90	84.9	203	5	US-10-501-838a-54 Sequence 54, Appl
39	85	80.2	23	4	US-10-665-184-29 Sequence 29, Appl
40	85	80.2	23	5	US-10-942-300-29 Sequence 29, Appl
41	85	80.2	23	5	US-10-501-838a-29 Sequence 29, Appl
42	85	80.2	29	4	US-10-665-184-33 Sequence 33, Appl
43	85	80.2	29	5	US-10-942-300-33 Sequence 33, Appl
44	85	80.2	30	5	US-10-501-838a-33 Sequence 33, Appl
45	82	77.4	23	4	US-10-665-184-1 Sequence 1, Appl

ALIGNMENTS

```

RESULT 1
US-10-665-184-27
; Sequence 27, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-665-184-27

Query Match      100.0%; Score 106; DB 4; Length 22;
Best Local Similarity 100.0%; Pred No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MYDITLALAGICQSARLVQOLA 22
Db      1 MYDITLALAGICQSARLVQOLA 22

RESULT 2
US-10-942-300-27
; Sequence 27, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

```

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 27
;; LENGTH: 22
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-10-942-300-27

Query Match 100.0%; Score 106; DB 5; Length 22;
Best Local Similarity 100.0%; Pred. No. 3,8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDITLALAGICQSARLVQQLA 22
Db 1 MYDITLALAGICQSARLVQQLA 22

RESULT 3
US-10-501-838A-27

;; Sequence 27, Application US/10501838A
;; Publication No. US20050215478A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel A.

;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501 NATL
;; CURRENT APPLICATION NUMBER: US/10/501,838A
;; PRIOR FILING DATE: 2004-07-19
;; PRIOR APPLICATION NUMBER: PCT/IB03/00968
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: US 60/355,396
;; NUMBER OF SEQ ID NOS: 72
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 27
;; LENGTH: 22
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-10-501-838A-27

Query Match 100.0%; Score 106; DB 5; Length 22;
Best Local Similarity 100.0%; Pred. No. 3,8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDITLALAGICQSARLVQQLA 22
Db 1 MYDITLALAGICQSARLVQQLA 22

RESULT 4
US-10-665-184-37

;; Sequence 37, Application US/10665184
;; Publication No. US20040146549A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel
;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501CIP
;; CURRENT APPLICATION NUMBER: US/10/665,184
;; PRIOR FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: PCT/IB03/00968
;; PRIOR FILING DATE: 2003-02-07
;; PRIOR APPLICATION NUMBER: 60/355,396
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 37
;; LENGTH: 29
;; TYPE: PRT
;; ORGANISM: Artificial

;; FEATURE:
;; OTHER INFORMATION: Penetrating peptide
;; FEATURE:
;; NAME/KEY: MOD_RES
;; LOCATION: (1)..(1)
;; OTHER INFORMATION: ACETYLTATION
;; FEATURE:
;; NAME/KEY: PEPTIDE
;; LOCATION: (29)..(29)
;; OTHER INFORMATION: wherein Xaa is Lysine-NH2
US-10-665-184-37

Query Match 100.0%; Score 106; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDITLALAGICQSARLVQQLA 22
Db 1 MYDITLALAGICQSARLVQQLA 22

RESULT 5
US-10-942-300-37

;; Sequence 37, Application US/10942300
;; Publication No. US20050136103A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel
;; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-503
;; CURRENT APPLICATION NUMBER: US/10/942,300
;; PRIOR FILING DATE: 2004-09-16
;; PRIOR APPLICATION NUMBER: 10/665,184
;; PRIOR FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: 10/664,989
;; PRIOR FILING DATE: 2003-09-17
;; PRIOR APPLICATION NUMBER: 60/503,615
;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 37
;; LENGTH: 29
;; TYPE: PRT
;; ORGANISM: Artificial

;; OTHER INFORMATION: Acylated Penetrating Peptide
;; FEATURE:
;; NAME/KEY: MISC FEATURE
;; LOCATION: (29)..(29)
;; OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated
;; OTHER INFORMATION: with a fatty acid
US-10-942-300-37

Query Match 100.0%; Score 106; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYDITLALAGICQSARLVQQLA 22
Db 1 MYDITLALAGICQSARLVQQLA 22

RESULT 6
US-10-501-838A-37

;; Sequence 37, Application US/10501838A
;; Publication No. US20050215478A1
;; GENERAL INFORMATION:
;; APPLICANT: Ben-Sasson, Shmuel A.
;; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
;; FILE REFERENCE: 24348-501 NATL


```
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 37
/ LENGTH: 29
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic: penetrating peptide
/ NAME/KEY: MISC FEATURE
/ LOCATION: (26)..(26)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ FEATURE:
/ OTHER INFORMATION: groups of the lysine residue
/ NAME/KEY: MISC FEATURE
/ LOCATION: (29)..(29)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ FEATURE:
/ OTHER INFORMATION: groups of the lysine residue
/ NAME/KEY: MISC FEATURE
/ LOCATION: (29)..(29)
/ OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
/ OTHER INFORMATION: peptide via the free amino groups of the lysine residue
/ US-10-501-838A-37
```

```
Query Match          100.0%; Score 106; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 YDITLALAGICQSARLVQOLA 22
Db 1 YDITLALAGICQSARLVQOLA 22
```

```
RESULT 7
US-10-665-184-3
/ Sequence 3, Application US/10665184
/ Publication No. US20040146549A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ APPLICANT: Cohen, Elnat
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-501CIP
/ CURRENT APPLICATION NUMBER: US/10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
/ US-10-665-184-3
```

```
Query Match          95.3%; Score 101; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YDITLALAGICQSARLVQOLA 22
Db 3 YDITLALAGICQSARLVQOLA 23
```

```
RESULT 8
US-10-665-184-26
/ Sequence 26, Application US/10665184
/ Publication No. US20040146549A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ APPLICANT: Cohen, Elnat
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-501CIP
/ CURRENT APPLICATION NUMBER: US/10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 26
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
/ US-10-665-184-26
```

```
Query Match          95.3%; Score 101; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YDITLALAGICQSARLVQOLA 22
Db 3 YDITLALAGICQSARLVQOLA 23
```

```
RESULT 9
US-10-942-300-3
/ Sequence 3, Application US/10942300
/ Publication No. US20050136103A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ APPLICANT: Cohen, Elnat
/ TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-503
/ CURRENT APPLICATION NUMBER: US/10/942,300
/ PRIOR FILING DATE: 2004-09-16
/ PRIOR APPLICATION NUMBER: 10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 10/664,989
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 60/503,615
/ PRIOR FILING DATE: 2003-09-17
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 23
/ TYPE: PRT
/ ORGANISM: Escherichia coli
/ US-10-942-300-3
```

```
Query Match          95.3%; Score 101; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YDITLALAGICQSARLVQOLA 22
Db 3 YDITLALAGICQSARLVQOLA 23
```

```
RESULT 10
US-10-942-300-26
/ Sequence 26, Application US/10942300
/ Publication No. US20050136103A1
/ GENERAL INFORMATION:
```

```

; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-942-300-26
```

```

Query Match          95.3%; Score 101; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YDITLALAGICQSARLVQQLA 22
Db 3 YDITLALAGICQSARLVQQLA 23
```

```

RESULT 11
US-10-501-838a-3
```

```

; Sequence 3, Application US/10501838A
; Publication No. US20050213478A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Elnat
; APPLICANT: Ben-Sasson, Shmuel A.
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-501-838a-3
```

```

Query Match          95.3%; Score 101; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YDITLALAGICQSARLVQQLA 22
Db 3 YDITLALAGICQSARLVQQLA 23
```

```

RESULT 12
US-10-501-838a-26
```

```

; Sequence 26, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501 NATL
```

```

; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-501-838a-26
```

```

Query Match          95.3%; Score 101; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YDITLALAGICQSARLVQQLA 22
Db 3 YDITLALAGICQSARLVQQLA 23
```

```

RESULT 13
US-10-665-184-25
```

```

; Sequence 25, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-665-184-25
```

```

Query Match          95.3%; Score 101; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2 YDITLALAGICQSARLVQQLA 22
Db 4 YDITLALAGICQSARLVQQLA 24
```

```

RESULT 14
US-10-942-300-25
```

```

; Sequence 25, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
```

; PRIOR FILING DATE: 2003-09-17
 ; NUMBER OF SEQ ID NOS: 65
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 25
 ; LENGTH: 24
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-942-300-25

Query Match 95.3%; Score 101; DB 5; Length 24;
 Best Local Similarity 100.0%; Pred. No. 2.7e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICOSARLVQQLA 22
 |||||
 DB 4 YDITLALAGICOSARLVQQLA 24

RESULT 15
 US-10-501-838A-25
 ; Sequence 25, Application US/10501838A
 ; Publication No. US20050215478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ben-Sasson, Shmuel A.
 ; APPLICANT: Cohen, Binat
 ; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
 ; FILE REFERENCE: 24348-501 NATL
 ; CURRENT APPLICATION NUMBER: US/10/501,838A
 ; CURRENT FILING DATE: 2004-07-19
 ; PRIOR APPLICATION NUMBER: PCT/IB03/00968
 ; PRIOR FILING DATE: 2003-02-07
 ; PRIOR APPLICATION NUMBER: US 60/355,396
 ; PRIOR FILING DATE: 2002-02-07
 ; NUMBER OF SEQ ID NOS: 72
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 25
 ; LENGTH: 24
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-501-838A-25

Query Match 95.3%; Score 101; DB 5; Length 24;
 Best Local Similarity 100.0%; Pred. No. 2.7e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YDITLALAGICOSARLVQQLA 22
 |||||
 DB 4 YDITLALAGICOSARLVQQLA 24

Search completed: January 23, 2006, 12:09:01
 Job time : 50.5327 secs

100-443886-100

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:32:56 ; Search time 12.5546 seconds
(without alignments)
176.269 Million cell updates/sec

Title: US-10-501-838a-28

Perfect score: 111

Sequence: 1 MRNLRTSLLAGLCTAAQWVFV 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	97.3	538	2 A41042	microbial metalloprotease
2	49.5	44.6	378	2 T25176	hypothetical prote
3	49	44.1	757	2 B84790	hypothetical prote
4	48	43.2	400	2 B96029	probable beta-keto
5	47	42.3	265	2 B87511	conserved hypotet
6	47	42.3	309	2 T29452	hypothetical prote
7	47	42.3	542	1 ORB80A	oligopeptide-bind
8	47	42.3	3839	2 T49799	related to TOM1 pr
9	46	41.4	121	2 AE3590	acetyl-CoA acetyl
10	46	41.4	232	2 B98150	superoxide dismuta
11	46	41.4	401	2 A13115	beta-ketoadipyl Co
12	46	41.4	417	2 B69804	multidrug resistanc
13	46	41.4	419	2 D98171	beta-ketoadipyl Co
14	45	40.5	472	2 P82143	hypothetical prote
15	44	39.6	580	2 C85919	hypothetical prote
16	44	39.6	580	2 A56382	purine permease, b
17	44	39.6	598	2 T51033	hypothetical prote
18	43.5	39.2	258	2 AD2183	hypothetical prote
19	43	38.7	198	2 A75270	conserved hypotet
20	43	38.7	205	2 I64155	hypothetical prote
21	43	38.7	308	1 H65050	probable ATP-bind
22	43	38.7	308	1 D91074	probable ATP-bind
23	43	38.7	345	2 T12367	NADH2 dehydrogenas
24	43	38.7	375	2 S76663	hypothetical prote
25	43	38.7	401	2 D83618	beta-ketoadipyl Co
26	43	38.7	557	1 S76051	hypothetical prote
27	43	38.7	582	2 AG0650	periplasmic oligop
28	43	38.7	4976	2 T14185	peptide synthetase
29	42	37.8	109	2 F97795	hypothetical prote

30	42	37.8	208	2 AH0199	conserved hypotet
31	42	37.8	211	2 D84960	hypothetical prote
32	42	37.8	213	2 S19211	ycfC protein - Bac
33	42	37.8	213	2 D90829	hypothetical prote
34	42	37.8	213	2 B85687	hypothetical prote
35	42	37.8	231	2 AB3138	superoxide dismuta
36	42	37.8	250	2 G83276	precorrin-3 methyl
37	42	37.8	342	2 S23438	hypothetical prote
38	42	37.8	343	2 F83126	ferric enterobacti
39	42	37.8	400	2 T35019	beta-ketoadipyl-Co
40	42	37.8	401	2 G83197	probable acyl-CoA
41	42	37.8	481	2 T03527	cobyrinic acid synth
42	42	37.8	483	2 S52226	cobD protein - Rho
43	42	37.8	571	2 S50331	NADH2 dehydrogenas
44	42	37.8	831	2 AD3337	potassium efflux b
45	42	37.8	855	2 G70821	hypothetical prote

ALIGNMENTS

RESULT 1

A41042
microbial metalloprotease (EC 3.4.24.-) nprB precursor, extracellular - Bacillus subti
N/Altername names: extracellular neutral proteinase B nprB
C/Species: Bacillus subtilis
C/Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 09-Jul-2004
C/Accession: A41042; H69666
R/Titan, L.; Wu, X.C.; Wong, S.L.
J. Bacteriol. 173, 6364-6372, 1991
A/Title: Cloning and expression of a novel protease gene encoding an extracellular neut
A/Reference number: A41042; M01D:92011407; PMID:1917867
A/Accession: A41042
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-538 <TRA>
A/Cross-references: UNIPROT:P39899; UNIPARC:UPI0000601B8; GB:M62845; NID:G143255; PIDN
R/Kuntel, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berle
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Enllich, S.D.; Emerson, P.T.; Enlian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galle
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koster, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, S.; Lardinois
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Mauesel
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portecelli
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, i
A/Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; M01D:98044033; PMID:9384377
A/Accession: H69666
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-538 <RUN>
A/Cross-references: UNIPARC:UPI0000601B8; GB:299109; GB:AL009126; NID:G2633260; PIDN:C
A/Experimental source: strain 168
C/Genetics:
C/Gene: nprB
C/Superfamily: zinc metalloendopeptidase, neutral protease type (elastase)
C/Keywords: hydrolase; metalloprotease

Query Match 97.3% Score 108; DB 2; Length 538;
Best Local Similarity 95.7% Pred. No. 2e+09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRNLRTSLLAGLCTAAQWVFV 23
Db 1 MRNLRTSLLAGLCTAAQWVFV 23

RESULT 2
T25176
C:Accession: T25176
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
R:Wilkerson, J.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25176
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-378 <MT>
A:Cross-references: UNIPROT:062381; UNIPARC:UPI0000061191; EMBL:Z81129; PIDN:CAB03410.1;
A:Experimental source: clone T23F1
R:Gardner, A.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z20017
A:Accession: T25333
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-378 <MT>
A:Cross-references: UNIPROT:062381; UNIPARC:UPI0000061191; EMBL:Z81129; PIDN:CAB03410.1;
A:Experimental source: clone T23F1
A:Gene: CRSP:T26H5.5
A:Map position: 5
A:Introns: 42/2; 103/2; 151/2; 311/2
Query Match
Best Local Similarity 44.6%; Score 49.5; DB 2; Length 378;
Matches 11; Conservative 5; Mismatches 7; Indels 1; Gaps 1;
QY 1 MNLTRTSLILA-GLCTAAQWVF 23
DB 60 MRLTINAFGLGICDLRMWFI 83
RESULT 3
B84790
C:Accession: B84790
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Wolfat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84790
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <STO>
A:Cross-references: UNIPROT:Q9ZU03; UNIPARC:UPI000000C710; GB:AE002093; NID:g4056478; PI
A:Gene: At2g37220
A:Map position: 2
Query Match
Best Local Similarity 44.1%; Score 49; DB 2; Length 757;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 3 NLTRTSLIAGLCTAAQWV 21
DB 359 NATYVSTLPGICDAGKWV 377
RESULT 4
B96029
C:Accession: B96029
C:Species: Streptomyces coelicolor
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
R:Parkhill, J.; Barrett, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, July 1998
A:Reference number: Z20619

C:Accession: B96029
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herman,
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: B96029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <KOR>
A:Cross-references: UNIPROT:Q92TL0; UNIPARC:UPI00000CB8B0; GB:AL591985; PIDN:CAC49901.1;
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Contents:
C:Genetics:
A:Gene: pcaf; SMD20589
A:Genome: Plasmid
C:Superfamily: acetyl-CoA acetyltransferase
C:Keywords: acyltransferase
Query Match
Best Local Similarity 43.2%; Score 48; DB 2; Length 400;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 RNLTRTSLIAGLCTAA 17
DB 65 RNLTARSLTNRGLTAA 18
RESULT 5
B87511
C:Accession: B87511
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Lamb, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n. J.; Kholmova, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87511
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <STO>
A:Cross-references: UNIPROT:Q9A6H6; UNIPARC:UPI00000C75F4; GB:AE005673; NID:g13423605; P
A:Gene: CC2118
Query Match
Best Local Similarity 42.3%; Score 47; DB 2; Length 265;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 MNLTRTSLIAGLCTAA 18
DB 1 MRLTARSLTNRGLTAA 18
RESULT 6
T29452
C:Accession: T29452
C:Species: Streptomyces coelicolor
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
R:Parkhill, J.; Barrett, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, July 1998
A:Reference number: Z20619

A/Accession: T29452
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-309 <PAR>
A/Cross-references: UNIPROT:086628; UNIPARC:UPI00000DAD26; EMBL:AL031155; PIDN:CAA20090.
C/Genetics:
A/Note: SC3A7.27
C/Superfamily: Streptomyces coelicolor probable integral membrane protein SC10A7.04

Query Match 42.3%; Score 47; DB 2; Length 309;
Best Local Similarity 66.7%; Pred. No. 7.2;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRRTSLLAGLCTAA 18
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 35 LRRTSLLAGLCTAA 49

RESULT 7
OREBOA
oligopeptide-binding protein precursor - *Salmonella typhimurium*
C/Species: *Salmonella typhimurium*
C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C/Accession: A25011
R/Hiles, I.D.; Higgins, C.F.
Eur. J. Biochem. 158, 561-567, 1986
A/Title: Peptide uptake by *Salmonella typhimurium*. The periplasmic oligopeptide-binding
A/Reference number: A25011; PMID:86274740; PMID:3525163
A/Residues: 1-542 <HIL>
A/Molecule type: DNA
A/Cross-references: UNIPROT:P06202; UNIPARC:UPI00001703A4; GB:X04194; NID:947807; PIDN:C
C/Comment: This protein binds peptides up to five amino acids long with high affinity.
C/Genetics:
A/Gene: oppA
A/Map position: 34 min
C/Superfamily: dipeptide transport protein
C/Keywords: binding protein-dependent transport system; oligopeptide transport; periplasmic
P/1-25/Domain: signal sequence #status predicted <SIG>
P/26-542/Product: oligopeptide-binding protein #status predicted <MAT>

Query Match 42.3%; Score 47; DB 1; Length 542;
Best Local Similarity 52.9%; Pred. No. 12;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MRNLTSLLAGLCTA 17
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 1 MSNLTSLLAGLCTA 17

RESULT 8
T49799
related to TOM1 protein [imported] - *Neurospora crassa*
N/Alternate names: protein B1822.10
C/Species: *Neurospora crassa*
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C/Accession: T49799
R/Schultze, U.; Nijn, V.; Hohnsbeil, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A/Reference number: 225022
A/Accession: T49799
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-383 <SCH>
A/Cross-references: UNIPROT:Q9P4Z1, UNIPARC:UPI000017B515; EMBL:AL356834; GSPDB:GN00116;
A/Experimental source: BAC clone B1822; strain OR14A
C/Genetics:
A/Gene: NCGP:B1822.10
A/Map position: 6
A/Intons: 16/3; 2607/1; 2658/1; 2845/1; 2987/2; 3204/3; 3694/1; 3809/1

Query Match 42.3%; Score 47; DB 2; Length 3839;
Best Local Similarity 52.6%; Pred. No. 74;

Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRNLTSLLAGLCTAAQ 19
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 679 LRNLTSLLAGLCTAA 697

RESULT 9
AE3590
acetyl-CoA acetyltransferase [imported] - *Brucella melitensis* (strain 16M)
C/Species: *Brucella melitensis*
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C/Accession: AE3590
R/Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Muir, C.; Log, T.; Ivanova,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A/Reference number: AD3252; PMID:11756688
A/Accession: AE3590
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-121 <KUR>
A/Cross-references: UNIPROT:Q8YC86; UNIPARC:UPI000058575; GB:AE008918; PIDN:AAU53888.1
A/Experimental source: strain 16M
C/Genetics:
A/Gene: EME110646
A/Map position: 11

Query Match 41.4%; Score 46; DB 2; Length 121;
Best Local Similarity 76.9%; Pred. No. 4.4;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RNLTRTSLLAGL 14
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 69 RNVARMSLLAGL 81

RESULT 10
B98150
superoxide dismutase (Fe), chloroplast [imported] - *Agrobacterium tumefaciens* (strain C
C/Species: *Agrobacterium tumefaciens*
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: B98150
R/Goodier, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tu*
A/Reference number: A97359; PMID:21608551; PMID:11743194
A/Accession: B98150
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-232 <KUR>
A/Cross-references: UNIPROT:Q8U6T0; UNIPARC:UPI00000D269D; GB:AE007870; PIDN:AAK88724.1
C/Genetics:
A/Gene: AGR_L_313
A/Map position: linear chromosome
C/Superfamily: superoxide dismutase (Mn)

Query Match 41.4%; Score 46; DB 2; Length 232;
Best Local Similarity 40.9%; Pred. No. 8;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 MRNLTSLLAGLCTAQMVF 22
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
Db 1 MNDITRTTILTAGAATATVAIF 22

RESULT 11
AI3115
beta-keoadipyl CoA thiolase [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AI3115

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Chen, L.; Wood, G.E.; Chen, Y.; Moe, I.
erage, G.; Gillet, W.; Grant, C.; Genthner, D.; Kutyavyn, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AF3115
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-401 <KUN>
A:Cross-references: UNIPROT:Q8U7A5; UNIPARC:UPI0000164954; GB:AE008689; PIDN:AAL45343.1;
A:Experimental source: strain C58 (Dugont)
C:Genetics:
A:Gene: pCAR
A:Map position: linear chromosome
C:Superfamily: acetyl-CoA acetyltransferase

Query Match 41.4%; Score 46; DB 2; Length 401;
Best Local Similarity 76.9%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RNLTRSLLAGL 14
DB 65 RNVARMSLLAGL 77

RESULT 12
B69804
multidrug resistance protein homolog yfiS - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
A:Accession: B69804
R:Kunst, F.; Ogasaara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hullo, M.F.
Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardin
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivolta, C.; Rooha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Ser
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Trepstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Wintere, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumestein, B.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: B69804
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-417 <KUN>
A:Cross-references: UNIPROT:Q31561; UNIPARC:UPI00000600DB; GB:Z99108; GB:AL009126; NID:9
A:Experimental source: strain 168
C:Genetics:
A:Gene: yfiS
C:Superfamily: tetracycline resistance protein

Query Match 41.4%; Score 46; DB 2; Length 417;
Best Local Similarity 55.0%; Pred. No. 14;
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 LTRISLLAGCTTAQMVFV 23
DB 214 LKRMPLLTGLTACVFLV 233

RESULT 13
D98171
beta-ketoadipyl-CoA thiolase (AY007371) [Imported] - *Agrobacterium tumefaciens* (strain C
C:Species: *Agrobacterium tumefaciens*

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
A:Accession: D98171
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Gurollo, B.; Goldman,
A.; Liu, F.; Mollat, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Matkeltz, B.;
Science 294, 2323-2328, 2001
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D98171
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <KUR>
A:Cross-references: UNIPROT:Q8U7A5; UNIPARC:UPI00000D25BE; GB:AE007870; PIDN:AAK88894.1;
C:Genetics:
A:Gene: AGR L 640
A:Map position: linear chromosome
C:Superfamily: acetyl-CoA acetyltransferase

Query Match 41.4%; Score 46; DB 2; Length 419;
Best Local Similarity 76.9%; Pred. No. 14;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RNLTRSLLAGL 14
DB 83 RNVARMSLLAGL 95

RESULT 14
F82143
hypothetical protein VC1896 [Imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
A:Accession: F82143
R:Heidelberg, J.F.; Bisen, J.A.; Nelson, W.C.; Claydon, R.A.; Gwin, M.L.; Dodson, R.J.;
charlson, D.; Ermolaeva, M.D.; Vamthayan, U.; Base, S.; Qin, H.; Dragol, I.; Sellere, P
1, R.R.; Mekalanos, U.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: F82143
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <HEI>
A:Cross-references: UNIPROT:Q9KQV2; UNIPARC:UPI0000C313E; GB:AE004265; GB:AE003852; NID
A:Experimental source: serogroup O1, strain N16961, biotype El Tor
C:Genetics:
A:Gene: VC1896
A:Map position: 1

Query Match 40.5%; Score 45; DB 2; Length 472;
Best Local Similarity 76.9%; Pred. No. 22;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RNLTRSLLAG 13
DB 16 IRSLLTSLLAG 28

RESULT 15
C85919
hypothetical protein gucQ [Imported] - *Escherichia coli* (strain O157:H7, substrain EDL93
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
A:Accession: C85919
R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glanier, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, B.J.; Davis, N.W.; Lim, A.; Dinalanca, E.; Potamouls, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-308 <STO>
A:Cross-references: UNIPROT:Q8X4S5; UNIPARC:UPI00001658B4; GB:AE005174; NID:G12517153; P

A:Experimental source: strain 0157:H7, substrain EDL933

C:Genetics:

A:Gene: gutQ

C:Superfamily: probable ATP-binding protein gutQ; CBS homology

Query Match 39.6%; Score 44; DB 2; Length 308;

Best Local Similarity 47.6%; Pred. No. 22;

Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKNLRTSLLAGLCTRAQNV 21

DB 209 MKELSRIGLVAVCDPQQOV 229

Search completed: January 23, 2006, 10:43:10
Job time : 12.5546 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:31:19 ; Search time 73.3188 Seconds
(without alignments)
221.323 Million cell updates/sec

Title: US-10-501-838a-28
Perfect score: 111
Sequence: 1 MRNLTRTSLLAGLCTPAQMVFV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	97.3	324	2	Q6DTN5_9BURK
2	108	97.3	341	2	Q4QZ38_BACSU
3	108	97.3	538	1	NPBB_BACSU
4	51.5	46.4	538	2	Q92P42_RHIME
5	50	45.0	440	2	Q4N8Y0_WMTC
6	49.5	44.6	378	2	Q62381_CAEBL
7	49	44.1	246	2	Q91970_ONCMY
8	49	44.1	246	2	Q91976_ONCMY
9	49	44.1	757	2	Q9ZUJ3_ARATH
10	48	43.2	297	2	Q4FXJ0_LEIMA
11	48	43.2	400	2	Q92T10_RHIME
12	48	43.2	656	2	Q4HAU0_9DETO
13	48	43.2	705	2	Q4FW19_LEIMA
14	47	42.3	56	2	Q6ACAI_LEIKX
15	47	42.3	265	2	Q9A6H6_CAUCR
16	47	42.3	309	2	Q86628_STRCO
17	47	42.3	400	2	Q4LL10_9BURK
18	47	42.3	413	2	Q7UST0_RHOB
19	47	42.3	543	1	OPPA_SALT
20	47	42.3	582	2	Q57NR3_SALCH
21	47	42.3	4065	1	TOM1_NEUCR
22	46	41.4	121	2	Q8YCG6_BRIME
23	46	41.4	203	1	Y1850_FASMU
24	46	41.4	232	2	Q7CVF7_AGRIS
25	46	41.4	355	2	Q81AB7_CAEBL
26	46	41.4	400	2	Q62B43_BURMA
27	46	41.4	400	2	Q62DB8_BURMA
28	46	41.4	400	2	Q8FW29_BRUSP
29	46	41.4	400	2	Q9F4K2_BRUSP
30	46	41.4	401	2	Q8U7A5_AGRIS
31	46	41.4	402	2	Q8POP9_XANAC

32	46	41.4	405	2	Q63011_BRUPS	Q63011 burkholderi
33	46	41.4	417	2	Q31561_BACSU	Q31561 bacillus su
34	46	41.4	418	2	Q966M9_CAEBL	Q966M9 caenorhabdi
35	46	41.4	419	2	Q7CV85_AGRIS	Q7CV85 agrobacteri
36	46	41.4	420	2	Q966N0_CAEBL	Q966N0 caenorhabdi
37	46	41.4	452	2	Q4NDE4_9MTC	Q4NDE4 arthropacte
38	46	41.4	543	2	Q650B2_MANSM	Q650B2 manheimia
39	46	41.4	605	2	Q5IAU5_BACFN	Q5IAU5 bacteroides
40	46	41.4	605	2	Q64R93_BACFR	Q64R93 bacteroides
41	46	41.4	1170	2	Q58M60_9CAUD	Q58M60 cyanophaga
42	46	41.4	1308	2	Q9GPM8_CARE	Q9GPM8 caenorhabdi
43	45.5	41.0	862	2	Q7Q058_ANOGA	Q7Q058 anopheles
44	45	40.5	192	2	Q881Z2_PSEPK	Q881Z2 pseudomonas
45	45	40.5	210	2	Q5ATK6_EMENTI	Q5ATK6 aspergillus

ALIGNMENTS

RESULT 1
Q6DTN5_9BURK PRELIMINARY; PRT; 324 AA.
AC Q6DTN5;
DT 25-OCT-2004 (T-EMBLrel. 28, Created)
DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
DE Organophosphorus insecticide hydrolase.
GN Name=opdb;
OS Burkholderia sp. FDS-1.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=265945;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FDS-1;
RA Zhang Z., Hong Q., Li S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY646835; AAT67170.1; -; Genomic_DNA.
DR SMR; Q6DTN5; 29-322.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR InterPro; IPR001279; Blactamase-like.
DR Pfam; PF00753; Lactamase_B; 1.
KW Hydrolase.
SQ
SEQUENCE 324 AA; 34330 MW; 291D4321AB983AF1 CRC64;
Query Match 97.3%; Score 108; DB 2; Length 324;
Best Local Similarity 95.7%; Pred. No. 1.8e-08;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRNLTRTSLLAGLCTPAQMVFV 23
Db 1 MRNLTRTSLLAGLCTPAQMVFV 23

RESULT 2
Q4QZ38_BACSU PRELIMINARY; PRT; 341 AA.
AC Q4QZ38;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Neutral protease B (Fragment).
GN Name=npb;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NNG1;
RA Tran T.T., Nguyen Thu H., Nguyen Ngoc Q., Pham Van T.,
RA Duong Hong Q., Nguyen Tien M., Dinh Duy K.;
RT "Cloning and sequencing of the gene encoding neutral protease from
Bacillus subtilis NNG1 isolated from cultivated soil in Vietnam.";

RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ973636; CAJ01440.1; -; Genomic_DNA.
 KM Protease.
 FT NON_TER 341 341
 SQ SEQUENCE 341 AA; 38244 MW; ASDBCDB814C98A7 CRC64;
 Query Match 97.3%; Score 108; DB 2; Length 341;
 Best Local Similarity 95.7%; Pred. No. 1.9e-08;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRLTRISLLAGLCTAAQMVV 23
 |||||
 Db 1 MRLTKTSLLAGLCTAAQMVV 23

RESULT 3
 NRRB_BACSU STANDARD; PRT; 538 AA.
 AC P39859;
 DT 01-SEP-1995 (Rel. 31, Last sequence update)
 DT 01-SEP-1995 (Rel. 31, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Neutral protease B precursor (EC 3.4.24.-).
 GN Name=nprb; OrderedLocustNames=BSU11100;
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CX NCBI_TaxId=1423;
 RN [1]
 RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 224-228.
 RX MEDLINE=92011407; PubMed=1917867;
 RA Tyran L., Wu X.C., Wong S.L.;
 RT "Cloning and expression of a novel protease gene encoding an
 RT extracellular neutral protease from *Bacillus subtilis*.";
 RL J. Bacteriol. 173:6364-6372(1991).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=168;
 RX MEDLINE=97177785; PubMed=9025291;
 RA Levine A., Vannier F., Roche B., Autret S., Mavel D., Seror S.J.;
 RT "A 10.3 kbp segment from nprb to argJ at the 102 degrees region of the
 RT *Bacillus subtilis* chromosome.";
 RL Microbiology 143:175-177(1997).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=168;
 RX MEDLINE=98015415; PubMed=9353931;
 RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
 RT "Sequencing of regions downstream of addA (98 degrees) and cItg (289
 RT degrees) in *Bacillus subtilis*.";
 RL Microbiology 143:3305-3308(1997).
 RN [4]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).
 RC STRAIN=168;
 RX MEDLINE=9804403; PubMed=9384377; DOI=10.1038/36786;
 RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bartero M.G., Besieres P., Bolotin A., Borchert S.,
 RA Borriss R., Boudier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.-K., Codan J.-J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Deulacq F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emmerson P.T.,
 RA Denton K.-D., Eyringston J., Fabret C., Ferrari B., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.-Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
 RA Gutlepp G., Guy B.J., Haga K., Hatach J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstegge S., Hosono S., Hullo M.-F., Itaya M.,
 RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
 RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
 RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
 RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Manuel C.,
 RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,
 RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogilwa A., Oudega B.,
 RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
 RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,

RA Rey M., Reynolds S., Rieger M., Rivoira C., Rocha E., Roche B.,
 RA Rose M., Sadie Y., Sato T., Scanlan B., Schleich S., Schroeder R.,
 RA Scoffone P., Sekiguchi J., Sekowska A., Seror S.J., Serro P.,
 RA Shin B.-S., Soldo B., Sorokin A., Taccioni E., Takagi T., Takahashi H.,
 RA Takemaru K., Takeuchi M., Yamakoshi A., Yanaka T., Terpstra F.,
 RA Tognoni A., Tosato V., Uchiyama S., Vandendol M., Vannier F.,
 RA Vassaretiti A., Viari A., Wambuit R., Wedler E., Wedler H.,
 RA Weitzegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
 RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E.,
 RA Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 CC -1- CATALYTIC ACTIVITY: Similar, but not identical, to that of
 CC thermolysin.
 CC -1- COFACTOR: Binds 4 calcium ions per subunit (Potential).
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (Potential).
 CC -1- BIOPHYSICOCHEMICAL PROPERTIES:
 CC pH dependence:
 CC Optimum pH is 6.6;
 CC Temperature dependence:
 CC Retains 65% activity at 65 degrees Celsius;
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the peptidase M4 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: M62845; AAA2626.1; -; Genomic DNA.
 CC EMBL: Z79580; CAB01832.1; -; Genomic DNA.
 CC EMBL: Y09476; CAB70628.1; -; Genomic DNA.
 CC EMBL: Z99109; CAB12950.1; -; Genomic DNA.
 CC PIR: A41042; A41042.
 CC HSSP: P05806; INPC.
 CC DR MEROPS: M04.012; -;
 CC Subtilist; BG10691; nprb.
 CC DR InterPro: IPR006025; Pept_M_Zn_BS.
 CC DR InterPro: IPR001570; Peptidase_M4.
 CC DR InterPro: IPR01096; Propap_M4_M36.
 CC DR InterPro: IPR005075; Propap_Pepsy.
 CC DR Pfam: PF07504; FFP; 1.
 CC DR Pfam: PF03413; Pepsy; 1.
 CC DR Pfam: PF01447; Peptidase_M4; 1.
 CC DR Pfam: PF02868; Peptidase_M4_C; 1.
 CC DR PRINTS: PR00730; THERMOLYSIN.
 CC DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 CC KW Calcium; Complete proteome; Direct protein sequencing; Hydrolase;
 CC Metal-binding; Metalloprotease; Protease; Signal; Zinc; Zymogen.
 CC FT SIGNAL 1 28
 CC FT PROPEP 29 223
 CC FT CHAIN 224 538
 CC FT ACT_SITE 370 370
 CC FT ACT_SITE 453 453
 CC FT METAL 365 365
 CC FT METAL 369 369
 CC FT METAL 373 373
 CC FT METAL 393 393
 CC FT METAL 404 404
 CC FT METAL 406 406
 CC FT METAL 407 407
 CC FT METAL 409 409
 CC FT METAL 412 412
 CC FT METAL 415 415
 CC FT METAL 416 416
 CC FT METAL 419 419
 CC FT METAL 422 422

SO SEQUENCE 538 AA; 5936 MW; 487D28BD8F64F14E CRC64;

Query Match 97.3%; Score 108; DB 1; Length 538;

Best Local Similarity 95.7%; Pred. No. 2.9e-08;

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNR/RTSLLAGLCTAQMVFV 23

DB 1 MNR/RTSLLAGLCTAQMVFV 23

RESULT 4

Q92P42 RHIME PRELIMINARY; PRT; 538 AA.

AC Q92P42;

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DE PUTATIVE PEPTIDE-BINDING PERIPLASMIC ABC TRANSPORTER PROTEIN.

GN OrderedAccession=AB01951; ORFNames=SMC04293;

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

NCBI_TaxID=382;

NCBI_TaxID=382;

NUCLEOTIDE SEQUENCE.

RT STRAIN=1021;

RA MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;

RA Boileau D., Barloy-Hubler P., Gouzy J., Bothe G., Anpe F., Bauric J.,

Godard T., Goffeau A., Kahn D., Kles B., Lelure V., Masny D.,

Pohl T., Portetelle D., Puehler A., Purnelle B., Rampeger U.,

Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.,

"Analysis of the chromosome sequence of the legume symbiont

Sinorhizobium meliloti strain 1021."

RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

DR EMBL; AL591789; CAC46530.1; -; Genomic_DNA.

DR HSSP; P23847; IDPP.

DR GO; GO:0005215; P:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000914; SBP_bac_5.

DR Pfam; PF00496; SBP_bac_5; 1.

DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.

DR Complete proteome.

SO SEQUENCE 538 AA; 59160 MW; 4516D405B34184A0 CRC64;

Query Match 46.4%; Score 51.5; DB 2; Length 538;

Best Local Similarity 65.0%; Pred. No. 22;

Matches 13; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

QY 1 MNR/RTSLLAGLCTAQMVFV 20

DB 1 MNR-----LTLAAGCTAALM 15

RESULT 5

QANEYO_9M1CC

ID QANEYO_9M1CC PRELIMINARY; PRT; 440 AA.

AC QANEYO;

DT 13-SEP-2005 (TrEMBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)

DE Histidinol dehydrogenase (EC 1.1.1.23).

GN ORFNames=ArchdRAFT_2126;

OS Archaeobacter sp. FB24.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Micrococcales; Micrococcaceae; Arthrobacter.

NCBI_TaxID=290399;

RESULT 6

Q62381 CAREL

ID Q62381 CAREL PRELIMINARY; PRT; 378 AA.

AC Q62381;

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DE Hypothetical protein sw-33.

GN Name=sw-33; ORFNames=T26H5.5;

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

NCBI_TaxID=6239;

NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA).

RT STRAIN=Br16C01 N2;

RA MEDLINE=99069613; PubMed=9851916;

RG The C. elegans sequencing consortium;

RT "Genome sequence of the nematode C. elegans: a platform for

investigating biology."

RT Science 282:2012-2018(1998).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

EMBL; Z82056; CAB04856.1; -; Genomic_DNA.

EMBL; Z81129; CAB04856.1; -; Genomic_DNA.

EMBL; Z82056; CAB0410.1; JOINED; Genomic_DNA.

PIR; T25176; T25176.

DR Ensembl; T26H5.5; Caenorhabditis elegans.

DR WormBase; WBGene00005780; T26H5.5.

DR WormPep; T26H5.5; CE18293.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0001584; F:rhodopsin-like receptor protein signaln. . .; IEA.

DR GO; GO:0007186; P:C-protein coupled receptor protein signaln. . .; IEA.

DR InterPro; IPR010726; DUF1300.

DR Pfam; PF06976; DUF1300; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS50262; G PROTEIN RECBP FL 2; 1.

DR Complete proteome; Hypothetical protein.

SO SEQUENCE 440 AA; 45957 MW; 4C6561B87B1DBB6A CRC64;

RA Hammon N., Israni S., Pitluck S., Richardson P.;

"Sequencing of the draft genome assembly of Arthrobacter sp. FB24."

RT Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.

RM (2)

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=FB24;

RG US DOE Joint Genome Institute (PGF-ORNL);

RA Larimer F., Land M.;

"Annotation of the draft genome assembly of Arthrobacter sp. FB24."

RT Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an

EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is

preliminary data.

CC -1- SIMILARITY: Belongs to the histidinol dehydrogenase family.

DR EMBL; AAHG01000007; EAL9592.1; -; Genomic_DNA.

DR InterPro; IPR001692; Histidinol_dh.

DR InterPro; IPR012131; Hscd1_dh_prok.

DR Pfam; PF00815; Histidinol_dh_1.

DR PIRSF; PIRSF000099; Histidinol_dh_1.

DR PRINTS; PR00083; H0LDHDKGNAB.

DR ProDom; PD002680; Histidinol_dh_1.

DR TIGRFAMs; TIGR00069; h1ad; 1.

KW Oxidoreductase.

SO SEQUENCE 440 AA; 45957 MW; 4C6561B87B1DBB6A CRC64;

Query Match 45.0%; Score 50; DB 2; Length 440;

Best Local Similarity 60.0%; Pred. No. 34;

Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 LTRTSLLAGLCTAQMVFV 23

DB 336 LTRNMLAGLCTAQMVFV 355

SQL	SEQUENCE	378 AA;	42911 MW;	6EBE4CA75455896AB	CRC64;
	Query Match		44.6%;	Score 49.5;	DB 2; Length 378;
	Best Local Similarity		45.8%;	Pred. No. 34;	
	Matches	11; Conservative	5; Mismatches	7; Indels	1; Gaps
OY		1 MENLRTRSLLA-GLCTAAQWVF 23			
		: : : : :			
DB		60 MRTLTFNAFLGIGCDLARMPF 83			
	RESULT 7				
ID	091970_ONCMY		PRT;	246 AA.	
AC	091970;				
DT	01-OCT-2000 (TREMBLrel. 15, Created)				
DR	01-OCT-2000 (TREMBLrel. 15, Last sequence update)				
DE	01-OCT-2003 (TREMBLrel. 25, Last annotation update)				
GN	Tumor necrosis factor alpha.				
OS	Name=tnfa;				
OC	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).				
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopeterygii; Teleostei; Euteleostei;				
CC	Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.				
OX	NCBI_TaxID=8022;				
RN	(1)				
RP	NUCLEOTIDE SEQUENCE.				
RA	Wang T., Laining K.J., Zou J., Holland J., Hirono I., Aoki T.,				
RA	Secombes C.J.;				
RL	Submitted (May-2000) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; AJ278085; CAB92316.1; -; Genomic_DNA.				
DR	HSP; F01375; 4TSV.				
DR	GO; GO:0016020; C:membrane; IEA.				
DR	GO; GO:0005164; P:tumor necrosis factor receptor binding; IEA.				
DR	GO; GO:0006955; P:immune response; IEA.				
DR	InterPro; IPRO06053; TNF abc.				
DR	InterPro; IPRO02960; TNF_beta.				
DR	InterPro; IPRO06052; TNF_family.				
DR	InterPro; IPRO03636; TNF_subf.				
DR	Pfam; PF00229; TNF; 1.				
DR	PRINTS; PR01234; TNECROSISFCT.				
DR	PRINTS; PR01236; TNPEPTA.				
DR	Prodom; PD002012; TNF_subf; 1.				
DR	SMART; SMO0207; TNF; 1.				
DR	PROSITE; PS50049; TNF_2; 1.				
SQ	SEQUENCE	246 AA; 27112 MW; 8CD9BDF6876CB	CRC64;		
	Query Match		44.1%;	Score 49;	DB 2; Length 246;
	Best Local Similarity		64.3%;	Pred. No. 28;	
	Matches	9; Conservative	3; Mismatches	2; Indels	0; Gaps
OY		9 LLLAGICTAAQWVF 22			
		: : : : :			
DB		41 LLIAGICAAALLP 54			
	RESULT 8				
ID	091976_ONCMY		PRT;	246 AA.	
AC	091976;				
DT	01-OCT-2000 (TREMBLrel. 15, Created)				
DR	01-OCT-2000 (TREMBLrel. 15, Last sequence update)				
DE	01-OCT-2003 (TREMBLrel. 25, Last annotation update)				
GN	Purative tumour necrosis factor alpha.				
OS	Name=TNF-alpha;				
OC	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).				
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopeterygii; Teleostei; Euteleostei;				
CC	Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.				
OX	NCBI_TaxID=8022;				
RN	(1)				
RP	NUCLEOTIDE SEQUENCE.				
RA	MEDLINE=21153236; PubMed=11231283;				

RA Laling K.J., Wang T., Zou J., Holland J., Hong S., Boile N., Hirose I.,
RA Aoki T., Secombe C.J.;
RT "Cloning and expression analysis of rainbow trout *Oncorhynchus mykiss*
RT tumour necrosis factor- α ,"
RL Eur. J. Biochem. 268:1315-1322 (2001).
DR EMBL: AJ277604; CAB89521.1; -, mRNA.
DR HSSP: P01375; 4TSV.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO: GO:0006955; P:immune response; IEA.
DR InterPro: IPR006053; TNF_abC.
DR InterPro: IPR002960; TNF_delta.
DR InterPro: IPR006052; TNF_family.
DR InterPro: IPR003636; TNF_subf.
DR Pfam: PF00229; TNF_1.
DR PRINTS: PR01234; TNECROSISFCT.
DR PRINTS: PR01236; TNFBETA.
DR ProDom: PD002012; TNF_subf; 1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00049; TNF_2; 1.
FT CHAIN 86 246 putative tumour necrosis factor alpha.
SQ SEQUENCE 246 AA; 27113 MW; DCD8971DF867C6C6 CRC64;

Query Match 44.1%; Score 49; DB 2; Length 246;
Best Local Similarity 64.3%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0

QY 9 LLNLGLCTPAQWVP 22
||:|||||:|:
DB 41 LLINGLCAAAALLF 54

RESULT 9
Q9ZUJ3 ARATH
ID Q9ZUJ3 ARATH PRELIMINARY; PRT; 757 AA.
AC Q9ZUJ3-
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE Hypothetical protein Atg37230.
GN Name=Atg37230;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Rounsley S.D., Yin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Romling C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RL [3]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Heuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carrinci P., Chen H., Cheuk R., Hayaishizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai I., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tridip M.G., Wu T., Shinzaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005896; AAC98044.1; -, Genomic_DNA.
DR EMBL: BT005725; AA064144.1; -, mRNA.
DR PIR: B84790; B84790.
DR GO: GO:0005488; F:binding; IEA.
DR InterPro: IPR002885; PPR.
DR InterPro: IPR01990; TPR-like_helical.
DR Pfam: PF01535; PPR; 14.
DR TIGRFAMs: TIGR00756; PPR; 13.

KM Hypothetical protein; Repeat.
SQ SEQUENCE 757 AA; 85705 MW; 67646FB9P972AE3 CRC64;

Query Match 44.1%; Score 49; DB 2; Length 757;
Best Local Similarity 57.9%; Pred. No. 75;
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 3 NLRTRSLLAGLCTAAQMV 21
DB 359 NATVTSTLPGLCDACKMV 377

RESULT 10

QAFKJ0 LEIMA PRELIMINARY; PRT; 297 AA.

AC QAFKJ0 LEIMA PRELIMINARY; PRT; 297 AA.
DT 13-SRP-2005 (TrEMBLrel. 31, Created)
DT 13-SRP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SRP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LMJ 0855;
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxId=5664;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Fieldlin.
RA Bertman M., Sisk B., Rajandream M.-A., Aert R., Anupama A.,
RA Apostolon Z., Attilpoe P., Bason N., Bauser C., Beck A., Beverly S.M.,
RA Blanchettin G., Borzym K., Bothe G., Brusch C.V., Collins M.,
RA Cadag E., Clarion L., Clayton C., Coulson R.M.R., Cronin A.,
RA Cruz A.K., Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoeft A.,
RA Pazolina G., Foerker N., Frasch A.C., Fraser N., Fuchs M., Gabell C.,
RA Godle A., Goffeau A., Harris D., Hertz-Poweler C., Hilbert H., Horn D.,
RA Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
RA Lord A., Louie T., Marra M., Maury D., Matthews K., Michael S.,
RA Mottam J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H.,
RA Oliver K., O'Neill S., Pentony M., Pohl T.M., Price C., Purnelle B.,
RA Quail M.A., Rabinowitsch E., Reinhardt R., Rieger M., Rinta J.,
RA Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
RA Schater M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
RA Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
RA Volckart G., Wambut R., Warren T., Wedler H., Woodward J., Zhou S.,
RA Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrett B.,
RA Myler P.J.;
RT "The genome of the kinetoplastid parasite, *Leishmania major*."
RL Science 0:0-0(2005).
DR EMBL/CP000080; AAZ09669.1; -; Genomic_DNA.
KM Hypothetical protein.
SQ SEQUENCE 297 AA; 31638 MW; B7C9C301F81DPAF CRC64;

Query Match 43.2%; Score 48; DB 2; Length 297;
Best Local Similarity 56.5%; Pred. No. 47;
Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 MRNLRTSLLAGLCTAAQMV 23
DB 43 MRNRAMVLSGLAGLCTAAGAVAV 65

RESULT 11

092TL0 RHIME PRELIMINARY; PRT; 400 AA.

AC 092TL0; DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative beta-ketoadipyl CoA thiolase protein (EC 2.3.1.-).
GN Name=pcar; OrderedLocNames=RH1501; ORFNames=SM050589;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxId=382;

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=1021;
RC MEDLINE-21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoefer F.U., Hernandez-Lucas I., Becker A., Cowle A., Gouzy J.,
RA Golding B., Puhler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
fixing endosymbiont *Sinorhizobium meliloti*."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL/AL591985; CAC43901.1; -; Genomic_DNA.
DR F1R; E96029; E96029.
DR HSP; P27796; IAFW.
DR GO; GO:0008415; P:acyltransferase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF02803; Thiolase_C; 1.
DR Pfam: PF00108; Thiolase_N; 1.
DR TIGRPFAM; TIGR01930; AcCoA-C-Actrans; 1.
DR PROSITE; PS00098; THIO_LAS_1; 1.
DR PROSITE; PS00737; THIO_LAS_2; 1.
DR Acyltransferase; Complete proteome; Plasmid; Transferase.
SQ SEQUENCE 400 AA; 41749 MW; 632C4B283C624B8B CRC64;

Query Match 43.2%; Score 48; DB 2; Length 400;
Best Local Similarity 68.8%; Pred. No. 61;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 RNTRSLLAGLCTA 17
DB 65 RNTRSLLAGLCTA 80

RESULT 12

Q4HAU0_9DEIO PRELIMINARY; PRT; 656 AA.

AC Q4HAU0; DT 13-SRP-2005 (TrEMBLrel. 31, Created)
DT 13-SRP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SRP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein precursor.
GN ORFNames=DgeodRAFT_1406;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OX NCBI_TaxId=319795;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pfitnick S., Richardson P.;
RT "Sequencing of the draft genome assembly of *Deinococcus geothermalis*
RT DSM 11300."
RT Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of *Deinococcus geothermalis*
RT DSM 11300."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL/AH801000003; EAL83300.1; -; Genomic_DNA.
KM Hypothetical protein; Signal.
FT SIGNAL 1 21 Potential.
SQ SEQUENCE 656 AA; 68960 MW; 8AF6CA5D4A07CCAA CRC64;

Query Match 43.2%; Score 48; DB 2; Length 656;
 Best Local Similarity 62.5%; Pred. No. 95;
 Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 LTRSLLAGLCTAAQ 19
 |||:|||||
 Db 6 LTLVTLVSLVSGLSQSQ 21

RESULT 13
 Q4FW19 LEIMA PRELIMINARY; PRT; 705 AA.
 ID Q4FW19;
 AC Q4FW19;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DE Hypothetical protein.
 GN ORFNames=LMJ_1210;
 OS Leishmania major.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 RX PubMed=16020728; DOI=10.1126/science.1112680;
 RA Bertman M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A., Apostolou Z., Attipoe P., Bason N., Bauser C., Beck A., Beverley S.M., Bianchetti G., Borzym K., Bothe G., Brusch C.V., Collins M., Cadey E., Clatoni L., Clayton C., Coulson R.M., Cronin A., Cruz A.K., Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoeft A., Favelina G., Foster N., Frasch A.C., Fraser A., Fuchs M., Gabel C., Goble A., Goffeau A., Harris D., Hertz-Powier C., Hilbert H., Horn D., Huang Y., Klages S., Knights A., Kube M., Laikre N., Litvin L., Lord A., Louie T., Marra M., Masuy D., Matthews K., Mitchell S., Mottam J.C., Muller-Auer S., Munden H., Nelson S., Norbertczak H., Oliver K., O'Neill S., Pentony M., Pohl T.M., Price C., Purnelle B., Quail M.A., Rabbinowitch E., Reinhardt R., Rieger M., Rinta J., Robben W., Robertson L., Ruiz J.C., Rutter S., Saunders D., Schaefer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S., Shin H., Sivan D., Squares R., Squares S., Tosato V., Vogt C., Voickaert G., Wambut R., Warren T., Wedler H., Woodward J., Zhou S., Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B., Myler P.J.;
 RA "The Genome of the Kinetoplastid Parasite, Leishmania major.";
 RL Science 309:436-442(2005).
 DR EMBL, CP000081; AA214504.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 705 AA; 77608 MW; 2B61B1FF22975002 CRC64;

Query Match 43.2%; Score 48; DB 2; Length 705;
 Best Local Similarity 57.9%; Pred. No. 1e+02;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 RNLTRSLLAGLCTAAQ 20
 |||:|||||
 Db 471 RPTRTVTLVSLVSGLSQ 489

RESULT 14
 Q6ACAL LEIXX PRELIMINARY; PRT; 56 AA.
 ID Q6ACAL LEIXX;
 AC Q6ACAL LEIXX;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Hypothetical protein.
 GN OrderedLocNames=lxk23430;
 OS Leifsonia xyl1 (subsp. xyl1).
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Micrococcineae; Microbacteriaceae; Leifsonia.
 OX NCBI_TaxID=59736;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CTCB07;
 RX PubMed=15305603;
 RA Monteiro-Vitorello C.B., Camargo L.B.A., Van Sluys M.A., Kitchajima J.P., Truffi D., do Amaral A.M., Harakava R., de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y., Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G., Almeida M.F. Jr., Carver H., Coutinho L.L., El-Dorri H.A., Petro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S., Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.B., Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F., Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A., Tsai S.M., Zanca A.S., Simpson A.J.G., Brumley S.M., Seubal J.C.;
 RA "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia xyl1 subsp. xyl1.";
 RL Mol. Plant Microbe Interact. 17:827-836(2004).
 DR EMBL, AE016822; AAT89992.1; -; Genomic_DNA.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 56 AA; 5969 MW; 8A5045417A9A386D CRC64;

Query Match 42.3%; Score 47; DB 2; Length 56;
 Best Local Similarity 52.9%; Pred. No. 16;
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RSLSLLAGLCTAAQ 22
 |||:|||||
 Db 12 RNSLMGCLTVASQIV 28

RESULT 15
 Q9A6H6 CAUCR PRELIMINARY; PRT; 265 AA.
 ID Q9A6H6;
 AC Q9A6H6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DR Hypothetical protein CC2118.
 GN OrderedLocNames=CC2118;
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
 RA Nierman W.C., Feldblyum T.V., Land M.T., Paulsen I.T., Nelson K.E., Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Pocioka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J., Berry K.J., Uetebach T.R., Tran K., Wolf A.M., Yamatchevan J.J., Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RA "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL, AE005884; AAK24089.1; -; Genomic_DNA.
 DR PIR, B87511; B87511.
 DR TIGR, CC2118; -.
 DR GO, GO:0005576; C:extracellular region; IEA.
 DR InterPro, IPR001283; Allrgn_Y5/Tpx1.
 DR Pfam, PF00188; SCP, 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 265 AA; 28656 MW; 0FB41288AC113969 CRC64;

Query Match 42.3%; Score 47; DB 2; Length 265;
 Best Local Similarity 61.1%; Pred. No. 61;
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 RNLTRSLLAGLCTAA 18
 |||:|||||
 Db 1 MRLTARSLTRAGLTLAA 18

Tue Jan 24 08:04:30 2006

us-10-501-838a-28.rup

Page 7

Search completed: January 23, 2006, 10:40:58
Job time : 74.318 secs

(C) 2007 by the author(s)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:35:50 ; Search time 21.0917 Seconds
(without alignments)
90.156 Million cell updates/sec

Title: US-10-501-838A-28
Perfect score: 111
Sequence: 1 MNLTFTSLILAGLCTAAQMFV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patente AA: *
1: /cgn2_6/prodata/1/1aa/5.COMB.pep: *
2: /cgn2_6/prodata/1/1aa/6.COMB.pep: *
3: /cgn2_6/prodata/1/1aa/H.COMB.pep: *
4: /cgn2_6/prodata/1/1aa/PCITUS.COMB.pep: *
5: /cgn2_6/prodata/1/1aa/RE.COMB.pep: *
6: /cgn2_6/prodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	43.2	429	US-09-922-501-10	Sequence 10, Appl
2	45.5	41.0	659	US-09-902-540-11293	Sequence 11293, A
3	45	40.5	171	US-09-724-864-59	Sequence 59, Appl
4	44.5	40.1	586	US-10-104-047-3411	Sequence 3411, Ap
5	44	39.6	217	US-09-489-039A-7769	Sequence 7769, Ap
6	44	39.6	425	US-09-328-352-6416	Sequence 6416, Ap
7	44	39.6	580	US-08-677-049-12	Sequence 12, Appl
8	43	38.7	114	US-09-902-540-10640	Sequence 10640, A
9	43	38.7	402	US-09-489-039A-10321	Sequence 10321, A
10	43	38.7	413	US-09-489-039A-13838	Sequence 13838, A
11	43	38.7	484	US-09-252-991A-13240	Sequence 13240, A
12	43	38.7	735	US-09-252-991A-32240	Sequence 32240, A
13	42	37.8	372	US-09-252-991A-18377	Sequence 18377, A
14	42	37.8	510	US-09-252-991A-22017	Sequence 22017, A
15	42	37.8	628	US-09-252-991A-32414	Sequence 32414, A
16	41.5	37.4	482	US-09-949-016-7966	Sequence 7966, Ap
17	41.5	37.4	806	US-09-902-540-12422	Sequence 12422, A
18	41	36.9	375	US-09-270-767-38715	Sequence 38715, A
19	41	36.9	376	US-09-270-767-53932	Sequence 53932, A
20	41	36.9	644	US-10-104-047-2945	Sequence 2945, Ap
21	41	36.9	1081	US-09-369-364A-17	Sequence 17, Appl
22	41	36.9	1104	US-09-981-953A-4	Sequence 4, Appl1
23	41	36.9	1146	US-09-824-734-2	Sequence 2, Appl1
24	40.5	36.5	216	US-09-543-681A-7993	Sequence 7993, Ap
25	40.5	36.5	271	US-08-701-191A-34	Sequence 34, Appl
26	40.5	36.5	271	US-09-664-526-34	Sequence 34, Appl
27	40.5	36.5	659	US-08-391-615-2	Sequence 2, Appl1

28	40.5	36.5	659	US-08-426-509A-8	Sequence 8, Appl1
29	40.5	36.5	659	US-08-232-545-8	Sequence 8, Appl1
30	40.5	36.5	659	US-09-949-016-5880	Sequence 5880, Ap
31	40.5	36.5	659	US-09-977-261-8	Sequence 8, Appl1
32	40.5	36.5	659	PCT-US95-05008-8	Sequence 8, Appl1
33	40.5	36.5	673	US-09-949-016-7226	Sequence 7226, Ap
34	40	36.0	26	US-09-348-578-1	Sequence 1, Appl1
35	40	36.0	26	US-09-699-684-1	Sequence 1, Appl1
36	40	36.0	27	US-09-348-578-2	Sequence 2, Appl1
37	40	36.0	27	US-09-699-684-2	Sequence 2, Appl1
38	40	36.0	32	US-09-348-578-7	Sequence 7, Appl1
39	40	36.0	32	US-09-699-684-7	Sequence 7, Appl1
40	40	36.0	203	US-09-328-714A-6	Sequence 6, Appl1
41	40	36.0	315	US-09-270-767-44966	Sequence 44966, A
42	40	36.0	380	US-09-489-039A-13215	Sequence 13215, A
43	40	36.0	381	US-09-216-393B-8	Sequence 8, Appl1
44	40	36.0	386	US-09-902-540-11003	Sequence 11003, A
45	40	36.0	389	US-09-336-643A-27	Sequence 27, Appl

ALIGNMENTS

```

RESULT 1
US-09-922-501-10
; Sequence 10, Application US/09922501
; Patent No. 6720168
; GENERAL INFORMATION:
; APPLICANT: Dartois, Veronique A.
; APPLICANT: Hoch, James A.
; APPLICANT: Valle, Fernando
; APPLICANT: Kumar, Manoj
; TITLE OF INVENTION: 2, 5-DKG PERMEASES
; FILE REFERENCE: P-SR 4877
; CURRENT APPLICATION NUMBER: US/09/922,501
; PRIOR FILING DATE: 2001-08-03
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 09/633,294
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 09/677,032
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Pantoea citrea
US-09-922-501-10

Query Match      43.2%; Score 48; DB 2; Length 429;
Best Local Similarity 61.1%; Pred. No. 8.3;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Cy      1 MNLTFTSLILAGLCTAA 18
       :|:|||||:|:|
Db      79 VRKLVFTSLILKGFCAAA 96

RESULT 2
US-09-902-540-11293
; Sequence 11293, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT FILING DATE: US/09/902,540
; CURRENT APPLICATION NUMBER: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825

```

```

; SEQ ID NO 11293
;
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
;
US-09-902-540-11293

```

Query Match	41.0%	Score 45.5;	DB 2;	Length 659;
Best Local Similarity	44.4%	Pred. No. 35;		
Matches 12;	Conservative 2;	Mismatches 6;	Indels 7;	Gaps 1;

```

QY      1 MRNLRTSLLAGL-----CTAAQM 20
          |||||
DB      1 MRRFVRTSLLAGLGLASGLMSCSDAML 27

```

RESULT 3
US-09-724-864-59
; Sequence 59, Application US/09724864

```

1  APPLICANT: Watson, James D.
2  APPLICANT: Harrison, James G.
3  TITLE OF INVENTION: Polynucleotides, polypeptides expressed
4  TITLE OF INVENTION: by the polynucleotides and methods for their use.
5  FILE REFERENCE: 11000.1050U1
6  CURRENT APPLICATION NUMBER: US/09/724,864
7  CURRENT FILING DATE: 2000-11-28
8  PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
9  PRIOR FILING DATE: 1999-12-23
10 NUMBER OF SEQ. ID NOS: 72
11 SOFTWARE: FastSeq for Windows Version 4.0
12 SEQ ID NO 59
13 LENGTH: 171
14 TYPE: PRT
15 ORGANISM: Mouse
16 IS-09-724-864-59

```

Query Match	40.5%	Score 45;	DB 2;	Length 171;
Best Local Similarity	45.0%	Pred. No. 9.1;		
Matches	9;	Conservative	3;	Mismatches 8;
				Indels 0;
				Gaps 0

QY 4 LTRISLLAGCTAAQMFV 23
|||:|||||:
Db 50 LMRLLVLASFCTAVALVI 69

```

RESULT 4
US-10-104-047-3411
, Sequence 3411, Application US/10104047
, Patent No. 6943241
, GENERAL INFORMATION:
, APPLICANT: HELIX RESEARCH INSTITUTE
, TITLE OF INVENTION: NO. 6943241el full length cDNA
, FILE REFERENCE: H1-A0105
, CURRENT APPLICATION NUMBER: US/10/104,047
, CURRENT FILING DATE: 2002-03-25
, PRIOR APPLICATION NUMBER:
, PRIOR FILING DATE:
, NUMBER OF SEQ ID NOS: 4096
, SOFTWARE: PatentIn Ver. 2.1
, SEQ ID NO 3411
, LENGTH: 566
, TYPE: PRT
, ORGANISM: Homo sapiens
US-10-104-047-3411

```

Query Match	40.1%	Score 44.5;	DB 2;	Length 586;
Best Local Similarity	43.5%	Pred. No. 46;		
Matches 10;	Conservative 6;	Mismatches 6;	Indels 1;	Gaps 1.

Oy 1 MRNLTRTSLLAGL-CTAAQMFV 22
 :|::|||:: |||::|::|::
Db 425 LRKLSRTGLVGIGSLCAGOLY 447

```

RESULT 5
US-09-489-039A-7789
; Sequence 7789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709, 20040001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7789
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

```

Query Match	39.6%	Score 44;	DB 2;	Length 217;
Best Local Similarity	45.0%;	Pred. No. 18;		
Matches 9; Conservative	5;	Mismatches	0;	Gaps 0;

Qy 2 RNLTRSLLAGLCTAQM V 21
:| :|||:| :| :|
Db 7 KNYDYITLALAGVCAQLV 26

```

RESULT 6
US-09-328-352-6416
; Sequence 6416, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6416
; LENGTH: 426
; TYPE: PRY
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6416

```

Query Match	39.6%	Score 44	DB 2	Length 426
Best Local Similarity	69.2%	Pred. No. 38		
Matches 9; Conservative 2		Mismatches 2	Indels 0	Gaps 0

```

QY      2 RNLTRSLLAGL 14
         ||:|:|||||
Db      89 RNVARNALLAGL 101

```

RESULT 7
 US-08-677-049-12
 ; Sequence 12, Application US/08677049
 ; Patent No. 5858707
 ; GENERAL INFORMATION:
 ; APPLICANT: Guimaraes, M. Jorge
 ; APPLICANT: Bazan, J. Fernando
 ; APPLICANT: McClanahan, Terrill K.
 ; APPLICANT: Zlotnik, Albert
 ; TITLE OF INVENTION: PURIFIED MAMMALIAN NUCLEOBASE PERMEASES;
 ; TITLE OF INVENTION: NUCLEIC ACIDS, ANTIBODIES
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: DNA Research Institute
 ; STREET: 901 California Avenue
 ;

CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,049
FILING DATE: 03-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,788
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: CHING, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0511
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 580 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Region
LOCATION: 207..236
OTHER INFORMATION: /note= "Encompasses TM 4 of Figure
OTHER INFORMATION: 4"
FEATURE:
NAME/KEY: Region
LOCATION: 437..499
OTHER INFORMATION: /note= "Encompasses TM 9 and TM 10
OTHER INFORMATION: of Figure 4"
US-08-677-049-12
Query Match 39.6%; Score 44; DB 1; Length 580;
Best Local Similarity 52.9%; Pred. No. 55;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 4 LRTSLLAGLCTAAQM 20
DB 110 LVSTSLVSGLSAVQM 126
RESULT 8
US-09-902-540-10640
Sequence 10640, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 10640
LENGTH: 114
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-10640

Query Match 38.7%; Score 43; DB 2; Length 114;
Best Local Similarity 33.3%; Pred. No. 12;
Matches 7; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
QY 1 RNLTRTSLLAGLCTAAQM 21
DB 4 LRPLATAVLAAGACAVARLI 24
RESULT 9
US-09-489-039A-10321
Sequence 10321, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10321
LENGTH: 402
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10321
Query Match 38.7%; Score 43; DB 2; Length 402;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 5 TRTSLLAGLCTAAQM 21
DB 65 TRSLLTGWCVLAIV 81
RESULT 10
US-09-489-039A-13838
Sequence 13838, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 13838
LENGTH: 413
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13838
Query Match 38.7%; Score 43; DB 2; Length 413;
Best Local Similarity 69.2%; Pred. No. 54;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 RNLTRTSLLAGL 14
DB 78 RNVARWALLLAGL 90
RESULT 11
US-09-266-965-17
Sequence 17, Application US/09266965
Patent No. 6495348
GENERAL INFORMATION:
APPLICANT: Sherman, D

APPLICANT: Mao, Y
APPLICANT: Varoglu, M
APPLICANT: He, M
APPLICANT: Sheldon, P
TITLE OF INVENTION: MlcMycIn biosynthetic gene cluster
FILE REFERENCE: 600.456US1
CURRENT APPLICATION NUMBER: US/09/266,965
CURRENT FILING DATE: 1999-03-12
EARLIER APPLICATION NUMBER: US 08/624,447
EARLIER FILING DATE: 1996-08-19
EARLIER APPLICATION NUMBER: PCT/US94/11279
EARLIER FILING DATE: 1994-10-06
EARLIER APPLICATION NUMBER: US 08/133,963
EARLIER FILING DATE: 1993-10-07
NUMBER OF SEQ ID NOS: 145
SOFTWARE: FASTSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 484
TYPE: PRT
ORGANISM: Streptomyces lavendulae
US-09-266-965-17

Query Match 38.7%; Score 43; DB 2; Length 484;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 RNLRTSLLAGCTAAQWVF 22
DB 348 RNLVIAAGLITTAAGFIW 364

RESULT 12
US-09-252-991A-32240
Sequence 32240, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32240
LENGTH: 735
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32240

Query Match 38.7%; Score 43; DB 2; Length 735;
Best Local Similarity 69.2%; Pred. No. 11e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RNLRTSLLAGL 14
DB 400 RNVARMLLAGL 412

RESULT 13
US-09-252-991A-18377
Sequence 18377, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18377
LENGTH: 372
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18377

Query Match 37.8%; Score 42; DB 2; Length 372;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 RNLRTSLLAGCTAAQ 19
DB 335 RNFRRLITVGEVLAAR 352

RESULT 14
US-09-252-991A-22017
Sequence 22017, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22017
LENGTH: 510
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22017

Query Match 37.8%; Score 42; DB 2; Length 510;
Best Local Similarity 69.2%; Pred. No. 1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RNLRTSLLAGL 14
DB 173 RNLARNMLLAGL 185

RESULT 15
US-09-252-991A-32414
Sequence 32414, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32414
LENGTH: 628
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32414

Query Match 37.8%; Score 42; DB 2; Length 628;
 Best Local Similarity 47.4%; Pred. No. 1.3e+02;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 5 TRTSLLAGLCTRAQMFV 23
 | : | | | | | : | :
 Db 416 TSLALAGLCTPALVLL 434

Search completed: January 23, 2006, 10:46:50
 Job time : 21.0917 secs

This Page Blank (verso)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:53:39 ; Search time 3.21397 Seconds
(without alignments)
72.521 Million cell updates/sec

Title: US-10-501-838A-28

Perfect score: 111

Sequence: 1 MRNITRTSLIAGLCTAAQWVFV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA_New:*

- 1: /cgn2_6/ptocdata/2/pubpsa/US08_NEW_PUB.Dep:*
- 2: /cgn2_6/ptocdata/2/pubpsa/US06_NEW_PUB.Dep:*
- 3: /cgn2_6/ptocdata/2/pubpsa/US07_NEW_PUB.Dep:*
- 4: /cgn2_6/ptocdata/2/pubpsa/PCT_NEW_PUB.Dep:*
- 5: /cgn2_6/ptocdata/2/pubpsa/US05_NEW_PUB.Dep:*
- 6: /cgn2_6/ptocdata/2/pubpsa/US10_NEW_PUB.Dep:*
- 7: /cgn2_6/ptocdata/2/pubpsa/US11_NEW_PUB.Dep:*
- 8: /cgn2_6/ptocdata/2/pubpsa/US60_NEW_PUB.Dep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	38.7	246	6	US-10-467-657-3824
2	41	36.9	608	7	US-11-109-157A-11
3	41	36.9	613	7	US-11-109-157A-42
4	41	36.9	1103	7	US-11-109-157A-9
5	40	36.0	312	6	US-10-858-730-117
6	39.5	35.6	182	6	US-10-821-234-1114
7	39	35.1	168	7	US-11-000-463-824
8	38.5	34.7	457	6	US-10-131-826A-236
9	38	34.2	188	6	US-10-945-853-2
10	38	34.2	218	5	US-09-978-360A-670
11	38	34.2	361	6	US-10-485-517-125
12	38	34.2	361	6	US-10-485-517-295
13	38	34.2	388	7	US-11-082-389-220
14	38	34.2	428	6	US-10-131-826A-330
15	38	34.2	898	7	US-11-145-631-11
16	37	33.3	118	6	US-10-467-657-1726
17	37	33.3	129	6	US-10-821-234-1667
18	37	33.3	191	6	US-10-467-657-7442
19	37	33.3	251	6	US-10-485-517-126
20	37	33.3	338	6	US-10-793-626-2868
21	37	33.3	409	6	US-10-821-234-1425
22	37	33.3	488	7	US-11-194-246-325
23	37	33.3	551	6	US-10-055-877-276
24	37	33.3	551	6	US-10-055-877-277
25	37	33.3	680	6	US-10-467-657-7612

26	37	33.3	716	6	US-10-467-657-8370	Sequence 8370, App
27	37	33.3	1049	7	US-11-137-465-42	Sequence 42, App1
28	37	33.3	1062	7	US-11-137-465-43	Sequence 43, App1
29	36	32.4	138	5	US-09-978-360A-483	Sequence 483, App
30	36	32.4	206	6	US-10-467-657-3004	Sequence 3004, App
31	36	32.4	253	7	US-11-080-991-42	Sequence 42, App
32	36	32.4	270	6	US-10-467-657-5806	Sequence 5806, App
33	36	32.4	271	7	US-11-019-711-26	Sequence 26, App1
34	36	32.4	307	7	US-11-019-711-87	Sequence 87, App1
35	36	32.4	307	7	US-11-019-711-88	Sequence 88, App1
36	36	32.4	449	6	US-10-821-234-1075	Sequence 1075, App
37	36	32.4	482	6	US-10-793-626-24	Sequence 24, App1
38	36	32.4	487	6	US-10-131-826A-528	Sequence 528, App
39	35	31.5	226	6	US-10-467-657-4756	Sequence 4756, App
40	35	31.5	368	6	US-10-995-561-753	Sequence 753, App
41	35	31.5	418	6	US-10-995-561-754	Sequence 754, App
42	35	31.5	418	6	US-10-995-561-755	Sequence 755, App
43	35	31.5	418	6	US-10-995-561-756	Sequence 756, App
44	35	31.5	418	6	US-10-995-561-757	Sequence 757, App
45	35	31.5	418	6	US-10-995-561-757	Sequence 757, App

ALIGNMENTS

```
RESULT 1
US-10-467-657-3824
; Sequence 3824, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: PONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 3824
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3824

Query Match      38.7%  Score 43;  DB 6;  Length 246;
Best Local Similarity 60.0%  Pred. No. 4;
Matches 9;  Conservative 2;  Mismatches 4;  Indels 0;  Gaps 0;

CY      6  RTSLIAGLCTAAQW 20
DB      59  QASFLIIGLCSAAVM 73

RESULT 2
US-11-109-157A-11
; Sequence 11, Application US/11109157A
; Publication No. US20050277175A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: TRUNCATED ADAMTS MOLECULES
; FILE REFERENCE: 01997.030500.
; CURRENT APPLICATION NUMBER: US/11/109,157A
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: 60/562,685
; PRIOR FILING DATE: 2004-04-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 11
```

LENGTH: 608
TYPE: PRT
ORGANISM: homo sapiens
US-11-109-157A-11

Query Match 36.9%; Score 41; DB 7; Length 608;
Best Local Similarity 81.8%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NLTRTSLLAG 13
|||:||||
Db 90 NLTRSSRLAG 100

RESULT 3
US-11-109-157A-42
Sequence 42, Application US/11109157A
Publication No. US20050277175A1
GENERAL INFORMATION:

APPLICANT: Wyeth
TITLE OF INVENTION: TRUNCATED ADAMTS MOLECULES
FILE REFERENCE: 01997.030500.
CURRENT APPLICATION NUMBER: US/11/109.157A
CURRENT FILING DATE: 2005-04-18
PRIOR APPLICATION NUMBER: 60/562,685
PRIOR FILING DATE: 2004-04-15
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.3
SEQ ID NO 42
LENGTH: 619
TYPE: PRT
ORGANISM: homo sapiens
US-11-109-157A-42

Query Match 36.9%; Score 41; DB 7; Length 619;
Best Local Similarity 81.8%; Pred. No. 26;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NLTRTSLLAG 13
|||:||||
Db 90 NLTRSSRLAG 100

RESULT 4
US-11-109-157A-9
Sequence 9, Application US/11109157A
Publication No. US20050277175A1
GENERAL INFORMATION:
APPLICANT: Wyeth
TITLE OF INVENTION: TRUNCATED ADAMTS MOLECULES
FILE REFERENCE: 01997.030500.
CURRENT APPLICATION NUMBER: US/11/109.157A
CURRENT FILING DATE: 2005-04-18
PRIOR APPLICATION NUMBER: 60/562,685
PRIOR FILING DATE: 2004-04-15
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn version 3.3
SEQ ID NO 9
LENGTH: 1103
TYPE: PRT
ORGANISM: homo sapiens
US-11-109-157A-9

Query Match 36.9%; Score 41; DB 7; Length 1103;
Best Local Similarity 81.8%; Pred. No. 48;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NLTRTSLLAG 13
|||:||||
Db 90 NLTRSSRLAG 100

RESULT 5

US-10-858-730-117
Sequence 117, Application US/10858730
Publication No. US20050255568A1
GENERAL INFORMATION:

APPLICANT: Bailey, Richard B.
APPLICANT: Blomquist, Paul
APPLICANT: Doten, Reed
APPLICANT: Driggers, Edward M.
APPLICANT: Madden, Kevin T.
APPLICANT: O'Leary, Jessica
APPLICANT: O'Toole, George
APPLICANT: Trueheart, Joshua
APPLICANT: Walbridge, Michael J.
APPLICANT: Yorgey, Peter S.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
FILE REFERENCE: 14184-030001
CURRENT APPLICATION NUMBER: US/10/858.730
CURRENT FILING DATE: 2004-06-01
PRIOR APPLICATION NUMBER: US 60/475,000
PRIOR FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: US 60/551,860
PRIOR FILING DATE: 2004-03-10
NUMBER OF SEQ ID NOS: 364
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 117
LENGTH: 312
TYPE: PRT
ORGANISM: Lactobacillus plantarum
US-10-858-730-117

Query Match 36.0%; Score 40; DB 6; Length 312;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 12; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

QY 2 NLTRTSLLAGCTAQMVF 22
|||:|||||
Db 73 NLTRIMYGAIGCT-AQTAF 93

RESULT 6
US-10-821-234-1114
Sequence 114, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmani, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821.234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: pt_seq_genes version 1.0
SEQ ID NO 1114
LENGTH: 182
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1114

Query Match 35.6%; Score 39.5; DB 6; Length 182;
Best Local Similarity 68.4%; Pred. No. 12;
Matches 13; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 MRNLTRTSLLAGCTAAQ 19
|||:|||||
Db 32 MRFLATFLLA-LSTAAQ 49

RESULT 7

```

US-11-000-463-824
/ Sequence 824, Application US/11000463
/ Publication No. US20050266423A1
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Qian, Xiaohong B.
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Zhang, Jie
/ APPLICANT: Zhou, Ping
/ APPLICANT: Cao, Yi-Cheng
/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
/ FILE REFERENCE: 785CIPACN
/ CURRENT APPLICATION NUMBER: US/11/000,463
/ PRIOR FILING DATE: 2004-11-29
/ PRIOR APPLICATION NUMBER: 10/291,265
/ PRIOR FILING DATE: 2002-11-08
/ PRIOR APPLICATION NUMBER: PCT/US01/02623
/ PRIOR FILING DATE: 2001-01-25
/ PRIOR APPLICATION NUMBER: 09/922,279
/ PRIOR FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: 09/491,404
/ PRIOR FILING DATE: 2000-01-25
/ PRIOR APPLICATION NUMBER: 09/517,746
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 09/631,451
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 09/533,870
/ PRIOR FILING DATE: 2000-09-15
/ NUMBER OF SEQ ID NOS: 944
/ SOFTWARE: PatSeq for Windows Version 3.0
/ SEQ ID NO 824
/ LENGTH: 168
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-000-463-824

```

```

Query Match      35.1%; Score 39; DB 7; Length 168;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 5 TRTSLLAGCT 16
DB 42 TVSLMLAGLCT 53

```

```

RESULT 8
US-10-131-826A-236
/ Sequence 236, Application US/10131826A
/ Publication No. US20050245730A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Geriltsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Goddard, Paul J.
/ APPLICANT: Guiney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

```

/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3330R1C128
/ CURRENT APPLICATION NUMBER: US/10/131,826A
/ PRIOR FILING DATE: 2002-04-24
/ PRIOR APPLICATION NUMBER: 60/049911
/ PRIOR FILING DATE: 1997-06-18
/ PRIOR APPLICATION NUMBER: 60/056974
/ PRIOR FILING DATE: 1997-08-26
/ PRIOR APPLICATION NUMBER: 60/059113
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059115
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059117
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059122
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059184
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059352
/ PRIOR FILING DATE: 1997-09-19
/ PRIOR APPLICATION NUMBER: 60/059588
/ PRIOR FILING DATE: 1997-09-19
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 236
/ LENGTH: 457
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-131-826A-236

```

```

Query Match      34.7%; Score 38.5; DB 6; Length 457;
Best Local Similarity 47.6%; Pred. No. 47;
Matches 10; Conservative 3; Mismatches 5; Indels 3; Gaps 1;
QY 1 MRLTRTSLL---LGLCTAA 18
DB 174 LRLMTYTVLNSPLTGKCTAA 194

```

```

RESULT 9
US-10-945-853-2
/ Sequence 2, Application US/10945853
/ Publication No. US20050255117A1
/ GENERAL INFORMATION:
/ APPLICANT: Biogen, Inc.
/ APPLICANT: Santicola-Nadel, Michele
/ APPLICANT: Adkins, Heather
/ APPLICANT: Miklasz, Steven Donald
/ APPLICANT: Rayhorn, Paul
/ APPLICANT: Schiffer, Susan Gail
/ APPLICANT: Williams, Kevin
/ TITLE OF INVENTION: Cytoto-Specific Antibodies
/ FILE REFERENCE: BGN117CPCCN
/ CURRENT APPLICATION NUMBER: US/10/945,853
/ PRIOR FILING DATE: 2004-09-20
/ PRIOR APPLICATION NUMBER: PCT/US02/31462
/ PRIOR FILING DATE: 2002-10-01
/ PRIOR APPLICATION NUMBER: PCT/US02/11950
/ PRIOR FILING DATE: 2002-04-17
/ PRIOR APPLICATION NUMBER: 60/286,782
/ PRIOR FILING DATE: 2001-04-26
/ PRIOR APPLICATION NUMBER: 60/293,020
/ PRIOR FILING DATE: 2001-05-17
/ PRIOR APPLICATION NUMBER: 60/301,091
/ PRIOR FILING DATE: 2001-06-26
/ PRIOR APPLICATION NUMBER: 60/367,002
/ PRIOR FILING DATE: 2002-03-22
/ NUMBER OF SEQ ID NOS: 9
/ SEQ ID NO 2
/ LENGTH: 188
/ TYPE: PRT

```

ORGANISM: Homo sapiens
US-10-945-853-2

Query Match 34.2%; Score 38; DB 6; Length 188;
Best Local Similarity 37.5%; Pred. No. 21;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 7 TSLIAGICTAAQWVF 22
Db 173 TTFMLAGICLSIQSY 188

RESULT 10
US-09-978-360A-670
; Sequence 670, Application US/09978360A
; Publication No. US2006009633A9
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste Dumas Milne
; APPLICANT: Duclet, Aymeric
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; APPLICANT: Clusel, Catherine
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: 56, US4, CIP
; CURRENT APPLICATION NUMBER: US/09/978,360A
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: US 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: US 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: US 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: US 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 60/099,273
; PRIOR FILING DATE: -09-04
; PRIOR APPLICATION NUMBER: US 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 09/215,435
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: PCT/IB98/02122
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: US 09/247,155
; PRIOR FILING DATE: 1999-02-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 810
; SOFTWARE: Patent.pm
; SEQ ID NO 670
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -50...-1
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (44, 47, 58, 91, 99, 100, 107, 113, 116, 132, 134, 135, 140, 159)
; OTHER INFORMATION: unknown
US-09-978-360A-670

Query Match 34.2%; Score 38; DB 5; Length 218;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 6 RTSLIAGICTAAQWVF 21
Db 38 RTALIAVCCGSASTV 53

RESULT 11
US-10-485-517-125

Sequence 125, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629MO
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 125
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-125

Query Match 34.2%; Score 38; DB 6; Length 361;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MRNLTSLIAGICTAA 18
Db 8 MRTIAKTSIALGLITGA 25

RESULT 12
US-10-485-517-295
; Sequence 295, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629MO
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 295
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-295

Query Match 34.2%; Score 38; DB 6; Length 361;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 MRNLTSLIAGICTAA 18
Db 8 MRTIAKTSIALGLITGA 25

RESULT 13
US-11-082-389-220
; Sequence 220, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroeger, Burkhard

APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
TITLE OF INVENTION: TRANSPORT
FILE REFERENCE: BGI-131CPCN
CURRENT APPLICATION NUMBER: US/11/082,389
CURRENT FILING DATE: 2005-03-16
PRIOR APPLICATION NUMBER: US 09/603024
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/143262
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: US 60/51281
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19930487.4
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19930489.0
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19931549.3
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931550.7
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932134.5
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 446
SEQ ID NO 220
LENGTH: 388
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-11-082-389-220

Query Match 34.2%; Score 38; DB 7; Length 388;
Best Local Similarity 43.8%; Pred. No. 47;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 NUTRTSLLAGLCTAA 18
DB 282 NSTRLALIVGVTSA 297

RESULT 14
US-10-131-826A-330
Sequence 330, Application US/10131826A
Publication No. US20050245730A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Denoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Geo, Wei-Qiang
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACTIS ENCODING THE SAME
FILE REFERENCE: P330R1C128
CURRENT APPLICATION NUMBER: US/10/131,826A
CURRENT FILING DATE: 2002-04-24

PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 330
LENGTH: 428
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-826A-330

Query Match 34.2%; Score 38; DB 6; Length 428;
Best Local Similarity 69.2%; Pred. No. 52;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 RNUTRTSLLAGL 14
DB 168 RNUTRTSLLAGL 180

RESULT 15
US-11-145-631-11
Sequence 11, Application US/11145631
Publication No. US20060003409A1
GENERAL INFORMATION:
APPLICANT: Berlin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-329001
CURRENT APPLICATION NUMBER: US/11/145,631
CURRENT FILING DATE: 2005-06-06
PRIOR APPLICATION NUMBER: US/09/841,739
PRIOR FILING DATE: 2001-08-29
PRIOR APPLICATION NUMBER: US 09/697,089
PRIOR FILING DATE: 2000-10-26
PRIOR APPLICATION NUMBER: US 60/161,822
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 898
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: majority sequence
US-11-145-631-11

Query Match 34.2%; Score 38; DB 7; Length 898;
Best Local Similarity 40.0%; Pred. No. 1,2e+02;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNUTRTSLLAGL 15
DB 189 LQNIQKTPLEVAIC 203

Search completed: January 23, 2006, 12:09:43

Job time : 3.21397 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 09:20:59 ; Search time 76.3319 Seconds
(without alignments)
132.392 Million cell updates/sec

Title: US-10-501-838a-28

Sequence: 1 MNLTSLTLLAGLCTAAQMPV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: Genesegp21.*
2: genesegp1908.*
3: genesegp2000.*
4: genesegp2001.*
5: genesegp2002.*
6: genesegp2003.*
7: genesegp2003b.*
8: genesegp2004.*
9: genesegp2005.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	100.0	23	6	ADBI6911 Bacillus
2	111	100.0	23	9	AEBO8245 Bacillus
3	111	100.0	30	6	ADBI6915 Escherich
4	111	100.0	30	9	AEBO8249 Penetrat
5	108	97.3	25	6	ADBI6897 Bacillus
6	108	97.3	25	9	AEBO8229 Bacillus
7	108	97.3	538	8	ADME6760 Bacillus
8	48	43.2	414	5	AAMS0973 Pantocoe c
9	48	43.2	414	5	AAO19981 Protein c
10	48	43.2	429	5	AAU77015 2.5-diket
11	47	42.3	272	4	ABG09325 Novel hum
12	46	41.4	203	9	AEBO8262 Pasteurel
13	46	41.4	398	9	ADN22235 Bacteri
14	46	41.4	400	8	ADN21915 Bacteri
15	46	41.4	400	8	ADN24993 Bacteri
16	46	41.4	401	8	ADN2438 Bacteri
17	46	41.4	401	8	ADN2438 Bacteri
18	46	41.4	401	8	ADN2438 Bacteri
19	46	41.4	401	8	ADN2438 Bacteri
20	46	41.4	401	8	ADN2438 Bacteri
21	45.5	41.0	659	9	ABM92094 M. xanthu
22	45	40.5	115	4	AAU66603 Propionib
23	45	40.5	115	4	ABM63122 Propionib
24	45	40.5	151	4	AAU45209 Propionib

25	45	40.5	151	6	ABM41728
26	45	40.5	155	5	ABR72375
27	45	40.5	171	4	AAE05363
28	45	40.5	240	8	ADN14294
29	45	40.5	244	8	ADH00896
30	44.5	40.1	138	5	ADG79516
31	44.5	40.1	273	5	ADG79425
32	44.5	40.1	426	7	ABM87132
33	44.5	40.1	586	7	ABM65257
34	44	39.6	217	7	ABO61272
35	44	39.6	393	6	ADN21708
36	44	39.6	426	6	ADN5129
37	44	39.6	455	4	AAU64449
38	44	39.6	455	6	ABM60968
39	44	39.6	499	6	ABM64775
40	43	38.7	95	4	AAO05653
41	43	38.7	109	8	AD128905
42	43	38.7	109	8	ADR44973
43	43	38.7	114	9	ABM91441
44	43	38.7	170	8	ADP58677
45	43	38.7	205	4	AAW50230

ALIGNMENTS

RESULT 1
ADBI6911 standard; peptide; 23 AA.

ADBI6911;

20-NOV-2003 (first entry)

Bacillus subtilis NprB penetrating peptide 33.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;
infertility; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiant; antiarteriosclerotic;
osteopathic; cytostatic; nootropic.

Bacillus subtilis.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-IB000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating
or preventing e.g. endocrine disorders.

Claim 2, Page 15; 60pp; English.

This invention relates to a novel peptide sequences capable of
translocating across a biological barrier. Furthermore, it refers to
methods that use these peptides to facilitate penetration of a
biologically active effector molecule such as a drug or other therapeutic
agent across biological barriers e.g. epithelial or endothelial cells
sealed by tight junctions. This peptide is derived from a bacterial
toxin, an integral membrane or extracellular protein and can comprise an
anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
or enzyme. The effector molecule, however, can comprise for example
insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from NprB of Bacillus
 CC subtilis and is penetrating peptide 33 of the invention.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 111; DB 6; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.3e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MRNLRTSLLAGCTAAQWVFV 23
 Db 1 MRNLRTSLLAGCTAAQWVFV 23
 RESULT 2
 AEB08245
 ID AEB08245 standard; peptide; 23 AA.
 XX
 AC AEB08245;
 DT 25-AUG-2005 (first entry)
 XX
 DE Bacillus subtilis NprB penetrating peptide 33, SEQ ID NO: 28.
 XX
 KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; anti anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antipneumatic; cyostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Bacillus subtilis.
 XX
 PN US2005136103-A1.
 XX
 PD 23-JUN-2005.
 XX
 PF 16-SEP-2004; 2004US-00942300.
 XX
 PR 17-SEP-2003; 2003US-00664989.
 XX
 PR 17-SEP-2003; 2003US-00665184.
 XX
 PR 17-SEP-2003; 2003US-0503615P.
 XX
 PA (BENS/) BEN-SASSON S A.
 PA (COHEN/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 DR WPI; 2005-444089/45.
 XX
 PT Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 PS Claim 46; SEQ ID NO 28; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for trans epithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Bacillus subtilis NprB
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 111; DB 9; Length 23;
 Best Local Similarity 100.0%; Pred. No. 2.3e-10;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MRNLRTSLLAGCTAAQWVFV 23
 Db 1 MRNLRTSLLAGCTAAQWVFV 23
 RESULT 3
 ADB16915
 ID ADB16915 standard; peptide; 30 AA.
 XX
 AC ADB16915;
 DT 20-NOV-2003 (first entry)
 XX
 DE Escherichia coli peptide 3 coupled to imaging compound linker, IBW-005.
 XX
 KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW endocrine disorder; obesity; neoplastic disease; neuroprotective;
 KW cardiac; antiarteriosclerotic; osteopathic; cyostatic; neurotropic;
 KW imaging linker; penetrating peptide; IBW-005.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetyl"
 FT Modified-site 30 /note= "C-terminal amide"
 FT
 XX
 PN WO2003066859-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 07-FEB-2003; 2003WO-IB000968.
 XX
 PF 07-FEB-2002; 2002US-0355396P.
 XX
 PR

XX 20-NOV-2003 (first entry)
 DT Bacillus subtilis NprB penetrating peptide 12.
 XX
 XX penetrating peptide; epithelial; endohepatic; tight junction; diabetes;
 XX infertility; hormone; vitamin deficiency; neurodegenerative;
 XX cardiovascular; haematological; endocrine disorder; obesity;
 XX osteopathic; cytostatic; nootropic.
 XX
 XX Bacillus subtilis.
 XX WO2003066859-A2.
 XX
 XX 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Ben-Sasson SA, Cohen E;
 XX
 XX WPI, 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX
 XX Claim 2, Page 14; 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dargargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, haematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX Accordingly, the peptide of this invention can be used in compositions
 XX that have neuroprotective, cardiac, antiatherosclerotic, osteopathic,
 XX cytostatic or nootropic activities. This peptide is from NprB of Bacillus
 XX subtilis and is penetrating peptide 12 of the invention.
 XX
 XX Sequence 25 AA;
 XX
 XX Query Match 97.3%; Score 108; DB 6; Length 25;
 XX Best Local Similarity 95.7%; Pred. No. 7.6e-10;
 XX Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 MRNLTSTSLLAGLCTAAQMVV 23
 XX Db 1 MRNLTSTSLLAGLCTAAQMVV 23
 XX
 XX RESULT 6
 XX AEB08229
 XX ID AEB08229 standard; peptide; 25 AA.
 XX
 XX AC AEB08229;
 XX
 XX 25-AUG-2005 (first entry)
 XX
 XX Bacillus subtilis NprB penetrating peptide 12, SEQ ID NO: 12.
 XX

KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW nootropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; arteriosclerosis;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virucide;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cycostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 XX Bacillus subtilis.
 XX
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHE/) COHEN E.
 XX Ben-Sasson SA, Cohen E;
 XX
 XX WPI, 2005-444089/45.
 XX
 XX Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 56; SEQ ID NO 12; 59pp; English.
 XX
 XX The present invention relates to a pharmaceutical composition of
 XX penetrating peptides for transepithelial delivery of effector. The
 XX invention comprises the effector sequentially coupled with a counter ion
 XX and at least one hydrophobic agent, where the effector is selectively
 XX encapsulated into a complex. The invention is useful for translocating
 XX effectors across a biological barrier such as epithelial cells and
 XX endothelial cells during treatment and prevention of disease or
 XX pathological conditions (including endocrine disorders, diabetes,
 XX infertility, hormone deficiencies, osteoporosis, ophthalmological
 XX disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 XX Parkinson's disease, multiple sclerosis, Huntington's disease,
 XX cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 XX coagulable states, coronary disease, cerebrovascular events, metabolic
 XX disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 XX hematological disorder, anemia of different entities, immunologic and
 XX rheumatologic disorders, autoimmune diseases, immune deficiencies,
 XX infectious diseases, viral infections, bacterial infections, fungal
 XX infections, parasitic infections, neoplastic diseases, multi-factorial
 XX disorders, impotence, chronic pain, depression, different fibrosis states
 XX and short stature) and for mucosal vaccination against anthrax and
 XX hepatitis B. The present sequence is the Bacillus subtilis NprB
 XX penetrating peptide. This sequence is used in the effective translocation
 XX of aminoglycoside antibiotics and antifungal agents across an epithelial
 XX barrier.
 XX
 XX Sequence 25 AA;
 XX

Query Match 97.3%; Score 108; DB 9; Length 25;
Best Local Similarity 95.7%; Pred. No. 7.6e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRNLTRTSLLAGLCTAAQWVFV 23
DB 1 MRNLTRTSLLAGLCTAAQWVFV 23

RESULT 7

ID ADM66760 standard; protein; 538 AA.

XX ADM66760;

XX 03-JUN-2004 (first entry)

DE Bacillus sp. thermolysin-like protease (TLP) precursor protein 9.

XX thermolysin-like protease; TLP; S1' site; gluten degradation; wheat;

KW baking industry; beer clarification; brewing; denaturing; skin dewooling;

KW leathers; protein hydrolyzate production; artificial sweetener; aspartame;

XX precursor; enzyme.

XX Bacillus sp.

XX MO2004011619-A2.

XX 05-FEB-2004.

XX 28-JUL-2003; 2003MO-US023726.

XX 26-JUL-2002; 2002US-0398656P.

XX (STRA-) STRATAGENE.

XX Clark DD, Braman JC;

XX WPI; 2004-143847/14.

XX New thermolysin-like protease with substrate specificity for a basic or

PT an acidic amino acid, useful in biological and biomedical research,

PT identifying therapeutic agents and diagnostic markers, or producing

PT artificial sweeteners.

XX disclosure; SEQ ID NO 53; 82pp; English.

XX The invention relates to a novel thermolysin-like protease (TLP)

CC comprising an S1' site and modified to have a substrate specificity for a

CC basic or an acidic amino acid. The thermolysin-like protease of the

CC invention may be useful in proteolytic applications, biological and

CC biomedical research, identifying therapeutic agents and diagnostic

CC markers, characterizing cells and organisms that have undergone genetic

CC modifications, identifying unknown illnesses, characterizing polypeptides

CC or identifying biological samples. The thermolysin-like protease may also

CC be useful in industrial processes, such as the degradation of gluten from

CC wheat within the baking industry, clarification of beer within the

CC brewing industry, denaturing or dewooling of skins within the leather

CC industry, preparation of protein hydrolyzates or production of artificial

CC sweeteners like aspartame. The current sequence is that of a TLP

CC precursor protein of the invention.

XX Sequence 538 AA;

QY Query Match 97.3%; Score 108; DB 8; Length 538;

DB Best Local Similarity 95.7%; Pred. No. 2e-08; 0; Indels 0; Gaps 0;

Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRNLTRTSLLAGLCTAAQWVFV 23

DB 1 MRNLTRTSLLAGLCTAAQWVFV 23

RESULT 8
AAM50973
ID AAM50973 standard; protein; 414 AA.

XX AAM50973;

XX 15-MAY-2002 (first entry)

DE Pantoea citrea permease B.

XX Permease B; 2-keto-L-gulonate acid; 2,5-diketoglutarate; transporter;

KW enzyme.

XX Pantoea citrea.

XX WO200212528-A2.

XX 14-FEB-2002.

XX 03-AUG-2001; 2001WO-US024327.

XX 04-AUG-2000; 2000US-00633294.

XX 29-SEP-2000; 2000US-00677032.

XX (GENV) GENENCOR INT. INC.

XX (MICR-) MICROGENOMICS INC.

XX Kumar M, Valle F, Darcois VA, Hoch JA;

XX WPI; 2002-227162/28.

XX N-PSDB; ABA91874.

XX Enhancing 2-keto-L-gulonate (2-KLG) acid production, by increasing

PT transport of 2,5-diketoglutarate (2,5-DKG) acid into a host cell that

PT converts 2,5-DKG to 2-KLG, and culturing the cell.

XX disclosure; Fig 1H; 56pp; English.

XX The present sequence is that of Pantoea citrea permease B, an enzyme that

CC acts as a 2,5-diketoglutarate (2,5-DKG) transporter. The invention

CC provides a method for enhancing a host cell's biosynthetic production of

CC 2-keto-L-gulonate acid (2-KLG). This involves selecting a host cell that

CC has an at least partially intracellular synthetic pathway which utilizes

CC 2,5-DKG to produce 2-KLG, increasing the transport of 2,5-DKG into the

CC host cell while maintaining the integrity of the host cell, culturing the

CC host cell to produce 2,5-DKG, and producing 2-KLG. Improving the

CC transport of 2,5-DKG into the host cell involves transforming the host

CC cell with DNA encoding a 2,5-DKG transporter, such as permease B. The

CC host cell is preferably a bacterium or yeast cell, especially Escherichia

CC coli, Pantoea or Klebsiella. 2-KLG is readily converted to L-ascorbic

CC acid. The method can also be used for the conversion of sorbose or

CC sorbitol to 2-KLG acid through sorbose, for the reduction of 5-keto-D-

CC gluconate (5-KDG) acid to L-idonic acid, and for the reduction of 5-KDG

CC acid to L-gulonate acid

XX Sequence 414 AA;

QY Query Match 43.2%; Score 48; DB 5; Length 414;

DB Best Local Similarity 61.1%; Pred. No. 56;

Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MRNLTRTSLLAGLCTAA 18

DB 64 VRKLVTSTLLMGFCAAA 81

RESULT 9

ID AAO19981 standard; protein; 414 AA.

XX AAO19981;

AC AAO19981;

DT 24-MAY-2002 (first entry)

XX DE Protein of *Pantoea citrea* Permease B.
 XX KM Biosynthesis; intracellular pathway; extracellular substrate; transport;
 XX KW host cell integrity; sorbose; sorbitol; sorbose; ascorbic acid; 5-KDG;
 XX KW gluconic acid; L-Idonic acid; 5-keto-D-gluconic acid; L-gluconic acid;
 XX KW recombinant microorganism; *Pantoea citrea* permease B.
 XX OS *Pantoea citrea*.
 XX PN WO200212481-A2.
 XX PD 14-FEB-2002.
 XX PF 03-AUG-2001; 2001WO-US024600.
 XX PR 04-AUG-2000; 2000US-00633294.
 XX PR 29-SEP-2000; 2000US-00677032.
 XX PA (GENEV) GENENCOR INT INC.
 XX PI Kumar M, Valle F;
 XX DR WPI; 2002-241757/29.
 XX DR N-PSDB; AAK98900.
 XX PT Enhancing host cell's biosynthetic production of desired compound that is
 PT derived from partially intracellular pathway by increasing transport of
 PT extracellular substrate used in the synthetic pathway.
 XX PS Disclosure; Fig 1H; 60pp; English.
 XX CC The invention relates to enhancing biosynthesis by host cell of a
 CC compound derived from an, at least, partially intracellular pathway of
 CC the host cell, comprising selecting a host cell that has an intracellular
 CC pathway using extracellular substrate, increasing transport of the
 CC extracellular substrate into the host cell while maintaining integrity of
 CC the host cell, where transport is a rate limiting step in biosynthesis by
 CC the host cell, culturing the host cell to produce an extracellular
 CC substrate, and producing a desired compound. Preferable host cell
 CC microorganisms include *Klebsiella* and *Pantoea*. The method of the
 CC invention is particularly useful in conjunction with ascorbic acid
 CC intermediate synthesis, for example the conversion of 2,5-diketogluconic
 CC acid (2,5-DKG) to 2-keto-L-gluconic acid (2-KLG); the conversion of
 CC sorbose or sorbitol to 2-KLG via sorbose; the reduction of 5-keto-D-
 CC gluconic acid (5-KDG) to L-Idonic acid; and the reduction of 5-keto-D-
 CC gluconic acid to L-gluconic acid. The advantage of the invention is the
 CC production of desired chemical compounds by recombinant microorganisms is
 CC increased by alleviating a limiting factor to a desired end product
 CC production. This sequence represents the protein of the substrate
 CC transporter *Pantoea citrea* Permease B of the invention
 XX SQ Sequence 414 AA;
 SQ Query Match 43.2%; Score 48; DB 5; Length 414;
 Best Local Similarity 61.1%; Pred. No. 56;
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MENTRTSLILAGLCTAA 18
 : |||||
 Db 64 VRKLVTSLILMGFCAA 81
 : |||||
 RESULT 10
 ID AAU77015 standard; protein; 429 AA.
 XX AC AAU77015;
 XX DT 21-MAY-2002 (first entry)
 XX DE 2,5-diketo-D-gluconic acid (DKG) permease prmb.
 XX KW

KW DKG; 2,5-diketo-D-gluconic acid permease; prmb; enzyme;
 KW 2-keto-L-gulonic acid; 2-KLG; ascorbic acid; 2-keto-reductase.
 XX OS *Pantoea citrea*.
 XX KW WO200212468-A2.
 XX PD 14-FEB-2002.
 XX PF 03-AUG-2001; 2001WO-US024507.
 XX PR 04-AUG-2000; 2000US-00633294.
 XX PR 29-SEP-2000; 2000US-00677032.
 XX PA (MICR-) MICROGENOMICS INC.
 XX PA (GENEV) GENENCOR INT INC.
 XX PI Darrois VA, Hoch JA, Valle F, Kumar M;
 XX DR WPI; 2002-217190/27.
 XX DR N-PSDB; ABK10185.
 XX PT Novel isolated polypeptide having 2,5-diketo-D-gluconic acid permease
 PT activity, useful for increasing 2-keto-L-gulonic acid bioproduction, and
 PT thus ascorbic acid production.
 XX PS Claim 44; Fig 1; 90pp; English.
 XX CC The invention relates to an isolated polypeptide which has 2,5-diketo-D-
 CC gluconic acid (DKG) permease activity namely PE1, PE6, PE1, prmb, prmb
 CC and viax2. Also included are the nucleic acids encoding the proteins, an
 CC antibody specific for the proteins, and a vector comprising the nucleic
 CC acid. The vector is useful for enhancing 2-keto-L-gulonic acid (2-KLG)
 CC production which involves expressing the polypeptide encoded by the
 CC catalyses the conversion of 2,5-DKG to 2-KLG, and enzymes that catalyse
 CC the conversion of glucose to 2,5-DKG. The bacterial cell is deficient in
 CC endogenous 2-keto-reductase activity and is of the genus *Pantoea*. The
 CC method further involves converting 2-KLG to ascorbic acid. The permease
 CC is useful for enhancing uptake of 2,5-DKG, and also to stop 2,5-DKG
 CC uptake, where the permease is added to a culture medium to compete with
 CC membrane bound permeases for 2,5-DKG. The permease nucleic acid is
 CC expressed in bacterial cells in order to enhance the rate of uptake of
 CC 2,5-DKG by the cells, and as probes or primers to identify and isolate
 CC 2,5-DKG permease homologues from additional species or as templates for
 CC production of mutant permeases. Expression of one or more 2,5-DKG
 CC permeases in bacterial cells significantly increases the production of 2-
 CC keto-L-gulonic acid bioproduction from D-glucose, which lowers the cost
 CC of commercial production of ascorbic acid. The present sequence
 CC represents a 2,5-DKG permease from *P. citrea*, prmb
 XX SQ Sequence 429 AA;
 SQ Query Match 43.2%; Score 48; DB 5; Length 429;
 Best Local Similarity 61.1%; Pred. No. 58;
 Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 MENTRTSLILAGLCTAA 18
 : |||||
 Db 79 VRKLVTSLILMGFCAA 96
 : |||||
 RESULT 11
 ID ABG09325 standard; protein; 272 AA.
 XX AC ABG09325;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #9316.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSR-) HYSRQ INC.
 PA Drmanac RT, Liu C, Tang YT,
 PI WPI; 2001-639362/73.
 DR N-PSDB; AAS73512.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 39684; 103bp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridization probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 272 AA,
 SO
 Query Match 42.3%; Score 47; DB 4; Length 272;
 Best Local Similarity 52.9%; Pred. No. 52;
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Oy 5 TRTSLLAGCTAAGV 21
 Db 175 TRTSLLAGCTSAKV 191
 RESULT 12
 AEB08262
 ID AEB08262 standard; protein; 203 AA.
 XX AEB08262;
 AC 25-AUG-2005 (first entry)
 XX
 DT Pasteurella multocida protein, SEQ ID NO: 60.
 XX
 DE Pharmaceutical, therapeutic; endocrine disease; endocrine-gen.; diabetes;
 XX anti-diabetic; endocrine disease; gastrointestinal disease;
 KM metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KM

XX degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KM neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KM neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KM dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KM anticonvulsant; genetic disorder; cardiovascular disease;
 KM cardiovascular-gen.; atherosclerosis; arteriosclerosis;
 KM coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KM nutritional disorder; vitamin deficiency; renal disease; nephrotoxic;
 KM gastrointestinal disease; hematological disease; antineutic; anemia;
 KM autoimmune disease; immunosuppressive; immune deficiency;
 KM immunostimulant; infectious disease; antimicrobial; infection;
 KM erectile dysfunction; andrology; major depressive disorder;
 KM antidepressant; psychiatric disorder; pain; analgesic;
 KM bacterial infection; antibacterial; viral infection; virocidic;
 KM fungal infection; fungicide; parasitic infection; antiparasitic;
 KM renal failure; antiinfectivity; antipneumatic; cytostatic;
 KM antiinflammatory; hepatotropic; hepatitis B virus infection.
 XX
 OS Pasteurella multocida.
 XX
 PN US2005136103-A1.
 XX 23-JUN-2005.
 XX 16-SEP-2004; 2004US-00942300.
 XX 17-SEP-2003; 2003US-0064989.
 PR 17-SEP-2003; 2003US-0065184.
 PR 17-SEP-2003; 2003US-0503615P.
 XX (BENS/) BEN-SASSON S A.
 PA (COHE/) COHEN E.
 PI Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.
 DR Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX Disclosure; SEQ ID NO 60; 59bp; English.
 PS The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transmembrane delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease,
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, multi-factorial
 CC infections, parasitic infections, neoplastic diseases, fungal
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Pasteurella multocida protein
 CC containing penetrating peptide at N-terminal end.
 XX
 SO Sequence 203 AA,
 Query Match 41.4%; Score 46; DB 9; Length 203;
 Best Local Similarity 47.6%; Pred. No. 54;
 Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 Oy 1 MRUTRSLLAGCTAAGV 21

DB | | : | | | : | | : | |
1 MANNYDITLALAGVCGAAKTV 21

RESULT 13
ID ADN22235 standard; protein; 398 AA.
XX
AC ADN22235;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #4888.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 4888; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQL Sequence 398 AA;
Query Match 41.4%; Score 46; DB 8; Length 398;
Best Local Similarity 76.9%; Pred. No. 1.1e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RNLTSTLLAGL 14
| | : | | | : | | : | |
DB 63 RNVARMSLLAGL 75

RESULT 14
ID ADN21915 standard; protein; 400 AA.
XX
AC ADN21915;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #4568.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 4568; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 400 AA;

Query Match 41.4%; Score 46; DB 8; Length 400;
 Best Local Similarity 64.7%; Pred. No. 1.1e+02;

Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 RNLTRTSLLAGLCTAA 18
 ||:|||||
 DB 65 RNVARMSALLAGLPTEA 81

RESULT 15

ADN24993
 ID ADN24993 standard; protein; 400 AA.

XX AC ADN24993;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #7646.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
 XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 XX pathogen tolerance; pest tolerance; plant disease resistance;
 XX cell cycle pathway modification; plant growth regulator;
 XX homologous recombination; seed oil yield; protein yield; carbohydrate;
 XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 XX bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.
 XX PA (HINK/) HINKLE G J.
 XX PA (SLAT/) SLATER S C.
 XX PA (CHEN/) CHEN X.
 XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI, 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.

XX PS Claim 1, SEQ ID NO 7646; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 400 AA;

Query Match 41.4%; Score 46; DB 8; Length 400;
 Best Local Similarity 76.9%; Pred. No. 1.1e+02;

Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 RNLTRTSLLAGL 14
 ||:|||||
 DB 65 RNVARMSALLAGL 77

Search completed: January 23, 2006, 10:28:35
 Job time : 77.3319 secs

THE 1990 BUNK (LUNCH)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 11:52:43 ; Search time 52.8297 Seconds
(without alignments)
181.907 Million cell updates/sec

Title: US-10-501-838a-28
Perfect score: 111
Sequence: 1 MRNLTRTSLLAGLCTAAQMV FV 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/prodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppaa/US10_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111	100.0	23	4	US-10-665-184-28
2	111	100.0	23	5	US-10-942-300-28
3	111	100.0	23	5	US-10-501-838a-28
4	111	100.0	30	4	US-10-665-184-32
5	111	100.0	30	5	US-10-942-300-32
6	111	100.0	30	5	US-10-501-838a-32
7	108	97.3	25	4	US-10-665-184-12
8	108	97.3	25	5	US-10-942-300-12
9	108	97.3	25	5	US-10-501-838a-12
10	98	88.3	23	5	US-10-501-838a-72
11	51	45.9	1059	5	US-10-732-923-12277
12	48	43.2	414	4	US-10-343-359-11
13	48	43.2	429	3	US-09-942-501-10
14	48	43.2	429	3	US-10-343-359-10
15	47	42.3	429	5	US-10-787-267a-10
16	47	42.3	272	5	US-10-450-763-39684
17	46	41.4	203	4	US-10-665-184-60
18	46	41.4	203	5	US-10-942-300-60
19	46	41.4	203	5	US-10-501-838a-54
20	46	41.4	398	4	US-10-369-493-4888
21	46	41.4	400	4	US-10-369-493-4568
22	46	41.4	400	4	US-10-369-493-7646
23	46	41.4	401	4	US-10-369-493-11471
24	46	41.4	401	4	US-10-369-493-14444
25	46	41.4	401	4	US-10-369-493-14563
26	46	41.4	401	4	US-10-369-493-15179
27	46	41.4	409	4	US-10-369-493-7325

28	45	40.5	155	3	US-09-866-050A-699	Sequence 699, Ap
29	44.5	40.1	517	4	US-10-437-963-142568	Sequence 142568,
30	44.5	40.1	586	4	US-10-104-047-3411	Sequence 3411, Ap
31	44	39.6	61	4	US-10-425-115-329077	Sequence 329077,
32	44	39.6	149	4	US-10-437-963-104656	Sequence 104656,
33	44	39.6	393	4	US-10-369-493-10741	Sequence 10741, A
34	44	39.6	538	4	US-10-156-761-12581	Sequence 12581, A
35	43	38.7	74	4	US-10-425-115-295361	Sequence 295361,
36	43	38.7	99	4	US-10-425-115-210003	Sequence 210003,
37	43	38.7	109	4	US-10-369-430A-10	Sequence 10, Appl
38	43	38.7	114	4	US-10-437-963-148268	Sequence 148268,
39	43	38.7	128	4	US-10-424-599-238195	Sequence 238195,
40	43	38.7	170	5	US-10-739-930-8754	Sequence 8754, Ap
41	43	38.7	196	4	US-10-437-963-143549	Sequence 143549,
42	43	38.7	205	4	US-10-665-184-59	Sequence 59, Appl
43	43	38.7	205	5	US-10-942-300-59	Sequence 59, Appl
44	43	38.7	205	5	US-10-501-838a-53	Sequence 53, Appl
45	43	38.7	272	6	US-11-097-143-26481	Sequence 26481, A

ALIGNMENTS

RESULT 1
US-10-665-184-28
; Sequence 28, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-665-184-28
Query Match 100.0%; Score 111; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MRNLTRTSLLAGLCTAAQMV FV 23
Db 1 MRNLTRTSLLAGLCTAAQMV FV 23
RESULT 2
US-10-942-300-28
; Sequence 28, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

;; PRIOR FILING DATE: 2003-09-17
;; NUMBER OF SEQ ID NOS: 65
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 28
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Bacillus subtilis
US-10-942-300-28

Query Match 100.0%; Score 111; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 4,3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLRTSLLAGLCTAAQWVFV 23
Db 1 MRNLRTSLLAGLCTAAQWVFV 23

RESULT 3

US-10-501-838A-28
; Sequence 28, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-501-838A-28

Query Match 100.0%; Score 111; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 4,3e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLRTSLLAGLCTAAQWVFV 23
Db 1 MRNLRTSLLAGLCTAAQWVFV 23

RESULT 4

US-10-665-184-32
; Sequence 32, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial

;; FEATURE:
;; OTHER INFORMATION: Penetrating peptide
;; FEATURE:
;; NAME/KEY: MOD_RES
;; LOCATION: (1)..(1)
;; OTHER INFORMATION: ACETYLATION
;; FEATURE:
;; NAME/KEY: PEPTIDE
;; LOCATION: (30)..(30)
;; OTHER INFORMATION: wherein Xaa is Lysine-NH2
US-10-665-184-32

Query Match 100.0%; Score 111; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 5,7e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLRTSLLAGLCTAAQWVFV 23
Db 1 MRNLRTSLLAGLCTAAQWVFV 23

RESULT 5

US-10-942-300-32
; Sequence 32, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial

;; FEATURE:
;; OTHER INFORMATION: Acylated Penetrating Peptide
;; FEATURE:
;; NAME/KEY: MISC FEATURE
;; LOCATION: (30)..(30)
;; OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated
US-10-942-300-32

Query Match 100.0%; Score 111; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 5,7e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRNLRTSLLAGLCTAAQWVFV 23
Db 1 MRNLRTSLLAGLCTAAQWVFV 23

RESULT 6

US-10-501-838A-32
; Sequence 32, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elnat
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL

```
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ CURRENT FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 32
/ LENGTH: 30
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic; penetrating peptide
/ FEATURE:
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (27)..(27)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ FEATURE:
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (30)..(30)
/ OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
/ OTHER INFORMATION: groups of the lysine residue
/ FEATURE:
/ NAME/KEY: MISC_FEATURE
/ LOCATION: (30)..(30)
/ OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
/ OTHER INFORMATION: peptide via the free amino groups of the lysine residue
/ US-10-501-838a-32
```

```
Query Match          100.0%; Score 111; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRNLTRTSLLAGLCTPAQMV 23
Db 1 MRNLTRTSLLAGLCTPAQMV 23
```

```
RESULT 7
US-10-665-184-12
/ Sequence 12, Application US/10665184
/ Publication No. US20040146549A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ APPLICANT: Cohen, Elnat
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-501CIP
/ CURRENT APPLICATION NUMBER: US/10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 12
/ LENGTH: 25
/ TYPE: PRT
/ ORGANISM: Bacillus subtilis
/ US-10-665-184-12
```

```
Query Match          97.3%; Score 108; DB 4; Length 25;
Best Local Similarity 95.7%; Pred. No. 1.4e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRNLTRTSLLAGLCTPAQMV 23
Db 1 MRNLTRTSLLAGLCTPAQMV 23
```

```
RESULT 8
US-10-942-300-12
/ Sequence 12, Application US/10942300
/ Publication No. US20050136103A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel
/ APPLICANT: Cohen, Elnat
/ TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
/ FILE REFERENCE: 24348-503
/ CURRENT APPLICATION NUMBER: US/10/942,300
/ PRIOR FILING DATE: 2004-09-16
/ PRIOR APPLICATION NUMBER: 10/665,184
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 10/664,989
/ PRIOR FILING DATE: 2003-09-17
/ PRIOR APPLICATION NUMBER: 60/503,615
/ PRIOR FILING DATE: 2003-09-17
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 12
/ LENGTH: 25
/ TYPE: PRT
/ ORGANISM: Bacillus subtilis
/ US-10-942-300-12
```

```
Query Match          97.3%; Score 108; DB 5; Length 25;
Best Local Similarity 95.7%; Pred. No. 1.4e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRNLTRTSLLAGLCTPAQMV 23
Db 1 MRNLTRTSLLAGLCTPAQMV 23
```

```
RESULT 9
US-10-501-838a-12
/ Sequence 12, Application US/10501838A
/ Publication No. US20050215478A1
/ GENERAL INFORMATION:
/ APPLICANT: Ben-Sasson, Shmuel A.
/ APPLICANT: Cohen, Elnat
/ TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
/ FILE REFERENCE: 24348-501 NATL
/ CURRENT APPLICATION NUMBER: US/10/501,838A
/ PRIOR FILING DATE: 2004-07-19
/ PRIOR APPLICATION NUMBER: PCT/IB03/00968
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: US 60/355,396
/ PRIOR FILING DATE: 2002-02-07
/ NUMBER OF SEQ ID NOS: 72
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 12
/ LENGTH: 25
/ TYPE: PRT
/ ORGANISM: Bacillus subtilis
/ US-10-501-838a-12
```

```
Query Match          97.3%; Score 108; DB 5; Length 25;
Best Local Similarity 95.7%; Pred. No. 1.4e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRNLTRTSLLAGLCTPAQMV 23
Db 1 MRNLTRTSLLAGLCTPAQMV 23
```

```
RESULT 10
US-10-501-838a-72
/ Sequence 72, Application US/10501838A
/ Publication No. US20050215478A1
/ GENERAL INFORMATION:
```

```

; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Eliaz
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; TITLE OF INVENTION: Biological Barrier
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 72
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-501-838A-72

Query Match          88.3%; Score 98; DB 5; Length 23;
Best Local Similarity 95.2%; Pred. No. 4,5e-08;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NLRTSLLAGLCTAQMVFV 23
Db 1 NLRTSLLAGLCTAQMVFV 21

RESULT 11
US-10-732-923-12277
; Sequence 12277, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgeton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/6310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 12277
; LENGTH: 1059
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-10-732-923-12277

Query Match          45.9%; Score 51; DB 5; Length 1059;
Best Local Similarity 47.8%; Pred. No. 54;
Matches 11; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MRNLRTSLLAGLCTAQMVFV 23
Db 172 VRNIRHTTLVAALTAATFILLV 194

RESULT 12
US-10-343-359-11
; Sequence 11, Application US/10343359
; Publication No. US20040029234A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Enhanced 2-Keto-L-Gluonic Acid
; TITLE OF INVENTION: Production
; FILE REFERENCE: GC687-PC1A
; CURRENT APPLICATION NUMBER: US/10/343,359
; CURRENT FILING DATE: 2003-01-20
; PRIOR APPLICATION NUMBER: US 09/633,294
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 09/677,032
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 11
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Pantoea citrea
US-10-343-359-11

Query Match          43.2%; Score 48; DB 4; Length 414;
Best Local Similarity 61.1%; Pred. No. 58;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MRNLRTSLLAGLCTAA 18
Db 64 VRKLVTSLLMGFCAA 81

RESULT 13
US-09-922-501-10
; Sequence 10, Application US/09922501
; Patent No. US20020120119A1
; GENERAL INFORMATION:
; APPLICANT: Dattois, Veronique A.
; APPLICANT: Hoch, James A.
; APPLICANT: Valle, Fernando
; APPLICANT: Kumar, Manoj
; TITLE OF INVENTION: 2, 5-DKG PERMEASES
; FILE REFERENCE: P-SR 4877
; CURRENT APPLICATION NUMBER: US/09/922,501
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/633,294
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 09/677,032
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Pantoea citrea
US-09-922-501-10

Query Match          43.2%; Score 48; DB 3; Length 429;
Best Local Similarity 61.1%; Pred. No. 60;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MRNLRTSLLAGLCTAA 18
Db 79 VRKLVTSLLMGFCAA 96

RESULT 14
US-10-343-369-10
; Sequence 10, Application US/10343369
; Publication No. US20040030113A1
; GENERAL INFORMATION:
; APPLICANT: Dattois, Veronique A.
; APPLICANT: Hoch, James A.
; APPLICANT: Valle, Fernando
; APPLICANT: Kumar, Manoj
; APPLICANT: Microgenomics, Inc.
; TITLE OF INVENTION: 2, 5-DKG PERMEASES
; FILE REFERENCE: GC687-PTC
; CURRENT APPLICATION NUMBER: US/10/343,369
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 09/633,294
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 09/677,032
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 429
; TYPE: PRT
```

ORGANISM: Pantoea citrea
US-10-343-369-10

Query Match	43.2%	Score 48	DB 4	Length 425
Best Local Similarity	61.1%	Pred. No. 60		
Matches 11, Conservative		1, Mismatches 6	Indels 0	Gaps 0

```

QY      1 MRNLTRTSLLAGLCTAA 18
      :| | | | | | | |
Db      79 VRKLVPTSLLMGFCAAA 96

```

RESULT 15
US-10-787

```

1  Sequence 10, Application US/10787267A
2  Publication No. US20050032088A1
3  GENERAL INFORMATION:
4  APPLICANT: Dattole, Veronique A.
5  APPLICANT: Hoch, James A.
6  APPLICANT: Valle, Fernando
7  APPLICANT: Kumar, Manoj
8  TITLE OR INVENTION: 2,5-PKG Permeases
9  FILE REFERENCE: CG687-3-01
10 CURRENT APPLICATION NUMBER: US/10/787,267A
11 CURRENT FILING DATE: 2004-02-25
12 PRIOR APPLICATION NUMBER: US 09/922,501
13 PRIOR FILING DATE: 2001-08-03
14 PRIOR APPLICATION NUMBER: US 60/325,774
15 PRIOR FILING DATE: 2000-08-04
16 PRIOR APPLICATION NUMBER: US 60/421,141
17 PRIOR FILING DATE: 2000-09-29
18 NUMBER OF SEQ ID NOS: 22
19 SOFTWARE: FastSeq for Windows Version 4.0
20 SEQ ID NO 10
21 LENGTH: 429
22 TYPE: prt
23 ORGANISM: Pantoea citrea
24 US-10-787-267A-10

```

Query Match	43.2%	Score 48	DB 5	Length 429
Best Local Similarity	61.1%	Pred. No. 60		
Matches 11	Conservative	1	Mismatches 6	Indels 0
			Gaps	0

```

OY      1 MRNLTRTSLLLAGLCTAA 18
        :| | | | | | | |
Db      79 VRKLVTSTLLMGFCAAA 96

```

Search completed: January 23, 2006, 12:09:01
Job time : 52.8297 secs

This Page Blank (Justo)

Page 1 of 1

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 10:43:16 ; Search time 39 Seconds
(without alignments)
56.743 Million cell updates/sec

Title: US-10-501-838a-29
Perfect score: 117
Sequence: 1 NYHDIYALAGVCSARLVHQLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	97.4	205	164155	hypothetical prote
2	98	83.8	208	2 AH0199	conserved hypochet
3	98	83.8	213	2 S19211	ycfc protein - Bac
4	98	83.8	213	2 D90829	hypothetical prote
5	98	83.8	213	2 B85687	hypothetical prote
6	95	81.2	215	2 AB0647	conserved hypochet
7	73	62.4	211	2 D84960	hypothetical prote
8	56	47.9	205	2 B82237	conserved hypochet
9	50.5	43.2	531	2 C95338	hypothetical prote
10	48	41.0	204	2 E82680	conserved hypochet
11	46	39.3	1265	2 T21782	hypothetical prote
12	45	38.5	433	2 T25946	hypothetical prote
13	45	38.5	947	2 AH0891	adenyl-transferease
14	45	38.5	1187	2 T18355	hypothetical prote
15	45	38.5	1243	2 T17350	vik protein Dic
16	44	37.6	72	2 T46623	hypothetical prote
17	44	37.6	300	2 D81399	malate dehydrogena
18	44	37.6	455	2 F82345	conserved hypochet
19	44	37.6	464	2 F64970	hypothetical prote
20	44	37.6	464	2 D90985	hypothetical prote
21	44	37.6	464	2 G85830	hypothetical prote
22	43	36.8	317	2 A44156	lipopolysaccharide
23	43	36.8	335	2 S70671	hypothetical prote
24	43	36.8	467	2 C83195	hypothetical prote
25	43	36.8	1857	2 T50513	malate dehydrogena
26	42.5	36.3	380	2 B86773	chitinase A (impor
27	42.5	36.3	1215	2 T13916	hypothetical prote
28	42	35.9	208	2 A98225	ribonuclease D (lm
29	42	35.9	208	2 AP3061	

30	42	35.9	238	2 AH2178	orotidine 5' monoph
31	42	35.9	429	2 T24922	hypothetical prote
32	42	35.9	457	2 TC6551	chitinase (BC 3.2.
33	42	35.9	517	2 T50801	argininosuccinate
34	42	35.9	2946	2 T00867	hypothetical prote
35	41.5	35.5	261	2 AC0164	probable transport
36	41	35.0	183	2 S24980	hydroxymethylbilan
37	41	35.0	261	2 B66036	probable phenylase
38	41	35.0	273	2 B83551	hypothetical prote
39	41	35.0	301	2 T35986	probable tRNA pseu
40	41	35.0	320	2 E87702	malate dehydrogena
41	41	35.0	342	2 AE0867	galactose operon r
42	41	35.0	373	2 F70781	probable c1ta prot
43	41	35.0	390	2 T30020	hypothetical prote
44	41	35.0	429	2 S09852	hypothetical prote
45	41	35.0	897	2 T02808	conserved hypochet

ALIGNMENTS

RESULT 1

164155
hypothetical protein HI0638 - Haemophilus influenzae (strain Rd KW20)
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #ext_change 09-Jul-2004
C/Accession: 164155
R/Flatschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, C.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995
A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A./Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A/Reference number: A64000; PMID:95350630; PMID:7542800
A/Accession: 164155
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-205 <TIGR>
A/Cross-references: UNIPROT:P44796; UNIPARC:UP1000013A674; GB:U32747; GB:U42023; NID:91
A/Note: best homolog was a hypothetical protein from Escherichia coli

Query Match 97.4% Score 114; DB 2; Length 205;
Best Local Similarity 95.7% Pred. No. 6.8e-11;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCSARLVHQLA 23
Db 3 NYHDIYALAGVCSARLVHQLA 25

RESULT 2

AH0199
conserved hypothetical protein YP01637 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #ext_change 09-Jul-2004
C/Accession: AH0199
R/Parhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tariga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.; 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; PMID:21470413; PMID:11586360
A/Accession: AH0199
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-208 <XLR>
A/Cross-references: UNIPROT:Q8ZF06; UNIPARC:UP100000CD828; GB:AL590842; PMID:12490459.1
C/Genetics:
A/Genes: YP01637

Query Match 83.8% Score 98; DB 2; Length 208;
Best Local Similarity 82.6% Pred. No. 2.7e-08;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 4 NYDITLALAGICQSARLVQOLA 26

RESULT 3
S19211
ycfc protein - Escherichia coli (strain K-12)
C|Species: Escherichia coli
C|Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C|Accession: S19211; A64858
R|Green, S.M.; Drabble, W.T.
submitted to the EMBL Data Library, May 1991
A|Description: Molecular analysis of the purB-phoP region of Escherichia coli K12.
A|Reference number: S19210
A|Accession: S19211
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-213 <GRB>
A|Cross-references: UNIPROT:P25746; UNIPARC:UPI000013A673; EMBL:X59307; NID:g42582; PIDN:AB0647
A|Experimental source: strain K-12
R|Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
-A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A|Title: The complete genome sequence of Escherichia coli K-12.
A|Reference number: A64720; MUID:97426617; PMID:9278503
A|Accession: A64858
A|Status: nucleic acid sequence not shown; translation not shown
A|Molecule type: DNA
A|Residues: 1-213 <BLAT>
A|Cross-references: UNIPARC:UPI000013A673; GB:AB000213; GB:U00096; NID:g1787371; PIDN:AA
A|Experimental source: strain K-12, substrain MG1655
C|Genetics:
A|Gene: yfcC
A|Start codon: GTG

Query Match 83.8%; Score 98; DB 2; Length 213;
Best Local Similarity 82.6%; Pred. No. 2.8e-08;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 4 NYDITLALAGICQSARLVQOLA 26

RESULT 4
D90829
hypothetical protein ECs1604 [imported] - Escherichia coli (strain O157:H7, substrain R1
C|Species: Escherichia coli
C|Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C|Accession: D90829
R|Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A|Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A|Reference number: A99629; MUID:21156231; PMID:11258796
A|Accession: D90829
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-213 <HAY>
A|Cross-references: UNIPROT:O8X736; UNIPARC:UPI000000D0AD2; GB:BA000007; PIDN:BA035027.1;
C|Genetics:
A|Gene: ECs1604

Query Match 83.8%; Score 98; DB 2; Length 213;
Best Local Similarity 82.6%; Pred. No. 2.8e-08;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 4 NYDITLALAGICQSARLVQOLA 26

RESULT 5
B85687
hypothetical protein yfcC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C|Species: Escherichia coli
C|Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C|Accession: B85687
R|Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hiller, D.; Grobeck, B.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A|Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A|Reference number: A85480; MUID:21074935; PMID:11206551
A|Accession: B85687
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-213 <STO>
A|Cross-references: UNIPROT:O8X736; UNIPARC:UPI000000D0AD2; GB:AB005174; NID:g12514786; P
A|Experimental source: strain O157:H7, substrain EDL933
C|Genetics:
A|Gene: yfcC

Query Match 83.8%; Score 98; DB 2; Length 213;
Best Local Similarity 82.6%; Pred. No. 2.8e-08;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 4 NYDITLALAGICQSARLVQOLA 26

RESULT 6
AB0647
conserved hypothetical protein STY1273 [imported] - Salmonella enterica subsp. enterica
C|Species: Salmonella enterica subsp. enterica serovar Typh
A|Note: This species has also been called Salmonella typh
C|Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
R|Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
'S.; Mole, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A|Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A|Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A|Reference number: AB0502; MUID:21534947; PMID:11677608
A|Accession: AB0647
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-215 <PAR>
A|Cross-references: UNIPARC:UPI0000059F77; GB:AL513382; PIDN:CAD08357.1; PID:g16502402;
C|Genetics:
A|Gene: STY1273

Query Match 81.2%; Score 95; DB 2; Length 215;
Best Local Similarity 78.3%; Pred. No. 8.7e-08;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NYHDIYALAGVCGSARLVQOLA 23
||:|||||:|||||:|||||
Db 6 NYDITLALAGICQSARLVQOLA 28

RESULT 7
D84960
hypothetical protein [imported] - Buchnera sp. (strain Aps)
C|Species: Buchnera sp.
C|Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C|Accession: D84960
R|Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A|Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A|Reference number: A84930; MUID:20445173; PMID:10993077
A|Accession: D84960

A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-211 <STO>
 A/Cross-references: UNIPARC:UPI00000584FF; GB:AP000398; GSPDB:GN00144
 A/Experimental source: strain APS
 C/Genetics:
 A/Gene: ynfC; BU262

Query Match 62.4%; Score 73; DB 2; Length 211;
 Best Local Similarity 71.4%; Pred. No. 0.00032;
 Matches 15; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 HDIVLALAGVCGSARLVHQLA 23
 Db 5 HLTITSLAGICQSAHLVQOLA 25

RESULT 8

conserved hypothetical protein VC1127 [imported] - *Vibrio cholerae* (strain N16961 serogroup C)
 C/Species: *Vibrio cholerae*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: B82237
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Charidson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000

A/Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A/Reference number: AB2035; MUID:20406833; PMID:10952301

A/Accession: B82237
 A/Status: preliminary
 A/Molecule type: DNA

A/Residues: 1-205 <HEI>

A/Cross-references: UNIPROT:Q9KX9; UNIPARC:UPI00000C2E84; GB:AE004193; GB:AE003852; NID
 A/Experimental source: serogroup O1; strain N16961; biotype El Tor
 C/Genetics:
 A/Gene: VC1127
 A/Map position: 1

Query Match 47.9%; Score 56; DB 2; Length 205;
 Best Local Similarity 47.6%; Pred. No. 0.18;
 Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 3 HDIVLALAGVCGSARLVHQLA 23
 Db 6 YDRTIAFAGICQAVLVQOVA 26

RESULT 9

conserved hypothetical protein Sma1131 [imported] - *Sinorhizobium meliloti* (strain 1021) magaplasma
 C/Species: *Sinorhizobium meliloti*
 C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C/Accession: C95338
 R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barlow-Hubler, F.; Bowe, J.; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.; Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A/Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*
 A/Reference number: A95262; MUID:21386509; PMID:11481432

A/Accession: C95338
 A/Status: preliminary
 A/Molecule type: DNA

A/Cross-references: UNIPROT:Q92281; UNIPARC:UPI00000CB152; GB:AE006469; PIDN:AAK65269.1;
 A/Experimental source: strain 1021, megaplasma pSymA

R/Gallbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barlow-Hubler, P.; J. Chalm, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federle, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kise, E.; Komp, C.; LeLaurie, H.; Vandenhol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Contents: annotation
 C/Genetics:
 A/Gene: Sma1131
 A/Genome: plasmid

Query Match 43.2%; Score 50.5; DB 2; Length 531;
 Best Local Similarity 42.9%; Pred. No. 3.5;
 Matches 9; Conservative 8; Mismatches 3; Indels 1; Gaps 1;

Qy 2 YHDIVLALAGVCGSARLVHQL 22
 Db 334 FH-IVTASGWCAGRIRHRL 353

RESULT 10

conserved hypothetical protein XF1439 [imported] - *Xylella fastidiosa* (strain 945c)
 C/Species: *Xylella fastidiosa*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: B82680
 R/Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

A/Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A/Reference number: AB2515; MUID:20365717; PMID:10910347

A/Note: for a complete list of authors see reference number A59328 below

A/Accession: B82680
 A/Status: preliminary
 A/Molecule type: DNA

A/Cross-references: UNIPROT:Q9PDE0; UNIPARC:UPI00000C2723; GB:AE003974; GB:AE003849; NID

A/Experimental source: strain 945c
 R/Simpson, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H.

Biologia, M.R.S.; Docena, C.; El-Dorri, H.; Facinca, A.P.; Ferreira, A.V.S.
 submitted to GenBank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Klajima, J.P.; Krieger, J.E.; Kuranae, E.B.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, J.

A/Authors: Martins, E.M.F.; Matukuna, A.Y.; Menck, C.F.M.; Miranda, R.C.; Palmieri, C.Y.;

P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.J.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawaiah

A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tenuhko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;

A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A/Gene: XF1439

RESULT 11

hypothetical protein F3582.9 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T21782
 R/Lennard, N.
 submitted to the EMBL Data Library, November 1996

A/Reference number: Z19471

A/Accession: T21782
 A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-1265 <WIS>
 A/Cross-references: UNIPROT:Q62231; UNIPARC:UPI000007E856; EMBL:Z81528; PIDN:CAE04288.1
 A/Experimental source: clone F3582
 C/Genetics:

A:Gene: CESP:P35E2.9
A:Map position: 1
A:Introns: 41/1; 180/1; 250/1; 317/1; 379/2; 456/1; 520/1; 638/1; 709/2; 780/1; 850/1; 9

Query Match
Best Local Similarity 39.3%; Score 46; DB 2; Length 1265;
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 NYHDIVALAGVCSARLVHQL 14
DB 529 NYDVICTLAGICR 541

RESULT 12
T25946
hypothetical protein ZC196.4 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
R:Murray, J.
Submitted to the EMBL Data Library, April 1997
A:Description: The sequence of *C. elegans* cosmid ZC196.
A:Reference number: Z20115
A:Accession: T25946
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-433 <MUR>
A:Cross-references: UNIPROT:O01628; UNIPARC:UPI0000081188; EMBL:U97007; PIDD:AA852296.1;
C:Genetics:
A:Experimental source: strain Bristol N2; clone ZC196
A:Map position: 5
A:Gene: CESP:ZC196.4
A:Introns: 10/3; 44/3; 68/3; 117/3; 170/3; 240/3; 303/3; 390/1

Query Match
Best Local Similarity 38.5%; Score 45; DB 2; Length 433;
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
Matches 7; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 NYHDIVALAGVCSARLVHQL 22
DB 295 NYDDPIMLKNICERANSIHPI 316

RESULT 13
AH0891
adenyl-transferase [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi (strain
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
Ch, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.D.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; MUID:21534547; PMID:11677608
A:Accession: AH0891
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-947 <PAR>
A:Cross-references: UNIPARC:UPI00005A4A4; GB:AL513382; PIDD:CAD07726.1; PID:G16504278;
C:Genetics:
A:Gene: STY3380

Query Match
Best Local Similarity 38.5%; Score 45; DB 2; Length 947;
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 ALAGVCSARLVHQL 22
DB 76 ALAGVCSARLVHQL 90

RESULT 14
T18355
hypothetical protein P3 - *Mycoplasma hyorhinis*
C:Species: *Mycoplasma hyorhinis*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIVALAGVCSARLVHQL 18
DB 39 NYHDIVALAGVCSARLVHQL 56

RESULT 15
T17390
vr1k protein - *Dichelobacter nodosus*
C:Species: *Dichelobacter nodosus*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
R:Billington, S.J.; Huggins, A.S.; Johansen, P.A.; Crellin, P.K.; Cheung, J.K.; Katz, M
Infect. Immun. 67, 1277-1286, 1999
A:Title: Complete nucleotide sequence of the 27-kilobase virulence related locus (vr1) o
A:Reference number: Z18734; MUID:99150261; PMID:10024571
A:Accession: T17390
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1243 <BIL>
A:Cross-references: UNIPROT:O32489; UNIPARC:UPI00000AF11A; EMBL:U20246; NID:G3493323; PI
A:Experimental source: strain A198
C:Superfamily: Dichelobacter nodosus vr1k protein

Query Match
Best Local Similarity 38.5%; Score 45; DB 2; Length 1243;
C:Species: *Dichelobacter nodosus*
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 NYHDIVALAGVCSARLVHQL 21
DB 668 NYHDIVALAGVCSARLVHQL 668

Search completed: January 23, 2006, 11:32:44
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 10:41:08 ; Search time 160 Seconds

(without alignments)
101.420 Million cell updates/sec

Title: US-10-501-838a-29

Perfect score: 117
Sequence: 1 NYHDIYALAGVCSARLVHQLA 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: 1: uniprot_05.80.*

2: uniprot_aprot.*
3: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114	97.4	205	1 Y638_HAEIN	P44796 haemophilus
2	114	97.4	205	2 Q4QMS9_HAE18	Q4QMS9 haemophilus
3	98	83.8	208	1 Y1637_YERPE	Q8ZFG6 yersinia pe
4	98	83.8	208	2 Q69Q03_YERPS	Q69Q03 yersinia ps
5	98	83.8	213	1 YCFC_ECO57	Q8X736 escherichia
6	98	83.8	213	1 YCFC_ECO16	Q8F1D7 escherichia
7	98	83.8	213	1 YCFC_ECOLI	P25746 escherichia
8	98	83.8	213	1 YCFC_SHIFL	Q831F8 shigella fl
9	95	81.2	213	1 YCFC_SALTI	Q8Z7H0 salmonella
10	95	81.2	213	1 YCFC_SALTY	Q8Z7H0 salmonella
11	95	81.2	213	2 Q5PMJ3_SALPA	Q5PMJ3 salmonella
12	92	78.6	204	2 Q6SVV5_MANSN	Q6SVV5 manheimia
13	91	77.8	215	2 Q57Q01_SALCH	Q57Q01 salmonella
14	89	76.1	212	2 Y1850_PASMU	Q9CJF8 pasteurella
15	89	76.1	212	2 Q6D4B8_BRWCT	Q6D4B8 erwania car
16	81	69.2	208	1 Y2805_PROIL	Q7N3B4 photorhabd
17	73	62.4	211	1 Y262_BUCAT	P57350 buchera ap
18	72	61.5	215	1 Y1650_HAEDU	Q7V1V5 haemophilus
19	71	60.7	217	1 Y243_BUCBP	Q8S4M4 buchera ap
20	56	47.9	205	1 Y1127_VIBCH	Q8K8X9 vibrio chol
21	56	47.9	205	1 Y1129_VIBPA	Q87GM5 vibrio para
22	55	47.0	205	1 Y1347_VIBVY	Q7ML55 vibrio vuln
23	55	47.0	205	1 Y2927_VIBVU	Q8D8P4 vibrio vuln
24	55	47.0	205	2 Q5E3W6_VIBFL	Q5E3W6 vibrio fisc
25	53	45.3	577	2 Q4UVH4_XANCP	Q4UVH4 xanthomonas
26	53	45.3	577	2 Q8P8K9_XANCP	Q8P8K9 xanthomonas
27	51	43.6	526	2 Q6AQL6_DESPS	Q6AQL6 desulfofocale
28	50.5	43.2	531	2 Q9Z281_RHIME	Q9Z281 rhizobium m
29	50	42.7	206	2 Q7NGC7_GLOVI	Q7NGC7 glaciebacter
30	50	42.7	207	2 Q60CA8_METCA	Q60CA8 methylcocc
31	49.5	42.3	367	2 Q5U8T4_TOBAC	Q5U8T4 nicotiana t

32	49.5	42.3	367	2 Q9ZRX7_TOBAC	Q9ZRX7 nicotiana t
33	49	41.9	245	2 Q7VXT7_BORPE	Q7VXT7 bordetella
34	49	41.9	245	2 Q7W6Y2_BORPA	Q7W6Y2 bordetella
35	49	41.9	245	2 Q7W1X7_BORBR	Q7W1X7 bordetella
36	49	41.9	245	2 Q4S165_TDRMG	Q4S165 tetradion n
37	49	41.9	390	2 Q9VGO1_DROME	Q9VGO1 drosophila
38	49	41.9	468	2 Q8KRT4_CHLTE	Q8KRT4 chlorobium
39	49	41.9	469	2 Q8KRT4_CHLTE	Q8KRT4 chlorobium
40	48	41.0	204	1 Y6E1_XYLPFA	Q9D6U0 xyella fas
41	48	41.0	204	1 Y6E1_XYLPFA	Q87DM1 xyella fas
42	48	41.0	446	2 Q4TBT1_TETNG	Q4TBT1 tetradion n
43	47	40.2	171	2 Q7RHR2_PLAYO	Q7RHR2 plasmodium
44	47	40.2	316	1 KPRS_RHOBA	Q7UPM4 rhodospirillum
45	47	40.2	334	2 Q8IB66_PLAF7	Q8IB66 plasmodium

ALIGNMENTS

RESULT 1
Y638_HAEIN STANDARD; PRT; 205 AA.
ID Y638_HAEIN STANDARD; PRT; 205 AA.
AC P44796;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein H10638.
GN OrderedCusNames=H10638;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=RD / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Pletschmann R.D., Adams M.D., White O., Clayton R.A., Kirness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G.G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uettermann T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Pine L.D., Pritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gehm C.J., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
RL Science 269:496-512(1995).
CC - SIMILARITY: Belongs to the UPF0274 family.
CC -
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL: U32747; AAC22298.1; -; Genomic DNA.
CC PIR: I64155; I64155.
CC TIGR: H10638; -.
CC HAMAP: MF_00695; -; 1.
CC InterPro: IPR007451; DUF489.
CC Pfam: PF04356; DUF489; 1.
CC Complete proteome; Hypothetical protein.
CC SEQUENCE 205 AA; 23235 MW; 81F31B0186BF82DA CRC64;
Query Match 97.4%; Score 114; DB 1; Length 205;
Best Local Similarity 95.7%; Pred. No. 9.6e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
RESULT 2
O4QMS9_HAE18 PRELIMINARY; PRT; 205 AA.
AC O4QMS9;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DE Predicted protein involved in purine metabolism.
OS Haemophilus influenzae (strain 86-028NP).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
CX NCBI_TaxID=281310;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
RA Harrison A., Dyer D.W., Gillespy A., Ray W.C., Mungur R., Carson M.B.,
RA Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,
RA Munson R.S. Jr.;
RT "Genomic sequence of an otitis media isolate of nontypeable
RT Haemophilus influenzae: comparative study with H. influenzae serotype
RT d, strain KW20."
RL J. Bacteriol. 187:4627-4636(2005).
DR EMBL; CP000057; AAX87668.1; -; Genomic_DNA.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome.
SQ SEQUENCE 205 AA; 23221 MW; AB9FAC75D46DF2 CRC64;

Query Match 97.4%; Score 114; DB 2; Length 205;
Best Local Similarity 95.7%; Pred. No. 9,6e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NYHDIYALAGVCSARLVHOLA 23
Db 3 NYHDIYALAGVCSARLVHOLA 25

RESULT 3
Y1637_YERPE STANDARD; PRT; 208 AA.
ID Y1637_YERPE
AC Q8ZFO6; Q8D0Q1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein YP01637/Y1798/YP1767.
GN OrderedLocustNames=YPO1637, Y1798, YP1767;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
CX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titchell R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Baeham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holtroyd S., Jagsels K., Karlyshev A.V.,
RA Leather S., Moulton S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RT Nature 411:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burtland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.U., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
```

```
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RA "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=91001 / Biovar Mediaevalis;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans."
RL DNA Res. 11:179-197(2004).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ414149; CAC90459.1; -; Genomic_DNA.
CC EMBL; AE013782; AAM65366.1; ALT INIT; Genomic_DNA.
CC EMBL; AE017133; AAS61994.1; -; Genomic_DNA.
CC PIR; AH0199; AH0199.
CC SMK; Q8ZFO6; 2-207.
CC DR HAMAP; MF_00695; -; 1.
CC DR InterPro; IPR007451; DUF489.
CC DR Pfam; PF04356; DUF489; 1.
CC KW Complete proteome; Hypothetical protein.
CC CONFLICT 1 M -> MY (in Ref. 2)
SQ SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE497723 CRC64;

Query Match 83.8%; Score 98; DB 1; Length 208;
Best Local Similarity 82.6%; Pred. No. 3,5e-07;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NYHDIYALAGVCSARLVHOLA 23
Db 4 NYHDIYALAGVCSARLVHOLA 26

RESULT 4
O66903_YERPS PRELIMINARY; PRT; 208 AA.
ID O66903_YERPS
AC O66903;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=YPT82431;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
CX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RX Chai N.P.S.G., Carniel E., Larimer F.W., Landerin J., Stoutland P.O.,
RA Regala M.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Himebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Franchet V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.U., Garcia B.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX916398; CAH21669.1; -; Genomic_DNA.
DR SMK; O66903; 2-207.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
```

```

RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltenwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogsh A., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,
RA Quill M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhl CT18.";
RL Nature 413:848-852(2001).
RN [12]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner P.R.;
RT "Comparative genome of Salmonella enterica serovar Typh strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
CC DR EMBL, AL627269; CAD08357.1; ALT INIT; Genomic DNA.
CC DR EMBL, AE016839; AA069312.1; ALT_INIT; Genomic DNA.
CC DR SMR, Q827H0; 2-213.
CC DR HAMAP, MF_00695; -; 1.
CC DR InterPro, IPR007451; DUF489.
CC DR Pfam, PF04356; DUF489; 1.
CC KW Complete proteome; Hypothetical protein.
CC SQ SEQUENCE 213 AA; 22942 MW; 63192D1C806BD58 CRC64;

Query Match      81.2%; Score 95; DB 1; Length 213;
Best Local Similarity 78.3%; Pred. No. 1.1e-06;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCGSARLVHQA 23
DB 4 NYHDIYALAGVCGSARLVHQA 26

RESULT 10
YCF3_SALTY STANDARD; PRT; 213 AA.
ID YCF3_SALTY STANDARD; PRT; 213 AA.
AC Q82P25;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycf3
GN Name=ycf3; OrderedLocustNames=SPR1617;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ty2 / SCS1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan B., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT Ty2.";
RL Nature 413:852-856(2001).
CC -1- SIMILARITY: Belongs to the UPF0274 family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
CC DR EMBL, AE008754; AAL20162.1; ALT_INIT; Genomic DNA.
CC DR SMR, Q82P25; 2-213.
CC DR StyGene, SG27272; ycf3.
CC DR HAMAP, MF_00695; -; 1.
CC DR InterPro, IPR007451; DUF489.
CC DR Pfam, PF04356; DUF489; 1.
CC KW Complete proteome; Hypothetical protein.
CC SQ SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match      81.2%; Score 95; DB 1; Length 213;
Best Local Similarity 78.3%; Pred. No. 1.1e-06;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCGSARLVHQA 23
DB 4 NYHDIYALAGVCGSARLVHQA 26

RESULT 11
QSPM33_SALPA PRELIMINARY; PRT; 213 AA.
ID QSPM33_SALPA PRELIMINARY; PRT; 213 AA.
AC QSPM33;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein ycf3
GN Name=ycf3; OrderedLocustNames=SPR1617;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.B., Clifton S.W., Latreille P.,
RA Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McClellan M.,
RA Harjane C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Kohlberg S., Strong C., Du F., Carter J., Kremzlik C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
RC EMBL, CP000026; AAV7544.1; -; Genomic DNA.
RC SMR, QSPM33; 2-213.
CC DR InterPro, IPR007451; DUF489.
CC DR Pfam, PF04356; DUF489; 1.
CC KW Complete proteome; Hypothetical protein.
CC SQ SEQUENCE 213 AA; 22916 MW; 63192D1C807C5D58 CRC64;

Query Match      81.2%; Score 95; DB 2; Length 213;
Best Local Similarity 78.3%; Pred. No. 1.1e-06;
Matches 18; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCGSARLVHQA 23
DB 4 NYHDIYALAGVCGSARLVHQA 26

RESULT 12
Q6SV55_MANSW PRELIMINARY; PRT; 204 AA.
ID Q6SV55;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.

```

```

GN OrderedLocusNames=M50298;
OS Mannheimia succiniciproducens (strain MBE155E);
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OX NCBI_TaxID=221988;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15378067; DOI=10.1038/nbt1010;
RA Hong S.H., Kim U.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
  Kim C.H., Jeong H., Hur C.G., Kim J.Y.;
RT "The genome sequence of the campylobilic rumen bacterium Mannheimia
  succiniciproducens."
RL Nac. Biotechnol. 22:1275-1281(2004).
DR EMBL; AF016827; A016905.1; -; Genomic_DNA.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 204 AA; 23036 MW; 3D64A70A30C33CB CRC64;

Query Match 78.6%; Score 92; DB 2; Length 204;
Best Local Similarity 78.3%; Pred. No. 3.1e-06;
Matches 18; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVHQLA 23
Db 4 NYDITLALAGVCGSARLVQCPA 26

RESULT 13
O57OC1 SALCH PRELIMINARY; PRT; 215 AA.
AC O57OC1
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Hypothetical protein y5cC.
GN Name=y5cC; OrderedLocusNames=SC1184;
OS Salmonella cholerae-suis (Salmonella enterica).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
  Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
  highly invasive and resistant zoonotic pathogen."
RL Nucleic Acids Res. 33:1690-1698(2005).
DR EMBL; AF017220; AA65090.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 215 AA; 23157 MW; C153B20629FAE8BD CRC64;

Query Match 77.8%; Score 91; DB 2; Length 215;
Best Local Similarity 77.3%; Pred. No. 4.8e-06;
Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVHQL 22
Db 6 NYDITLALAGVCGSARLVQOL 27

RESULT 14
Y1850 PASMU STANDARD; PRT; 203 AA.
AC Q9CJY8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein PM1850.
GN OrderedLocusNames=PM1850;
OS Pasteurella multocida.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L., Paustian M.L., Whitlam T.S., Kapur V.,
  "Complete genomic sequence of Pasteurella multocida Pm70."
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1 SIMILARITY: Belongs to the UPF0274 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AE006223; AA03934.1; -; Genomic_DNA.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 203 AA; 22717 MW; 5EB91EBF0E18EDF CRC64;

Query Match 76.1%; Score 89; DB 1; Length 203;
Best Local Similarity 73.9%; Pred. No. 9.4e-06;
Matches 17; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVHQLA 23
Db 3 NYDITLALAGVCGSARLVQCPA 25

RESULT 15
O6D4E8 ERWCT PRELIMINARY; PRT; 212 AA.
AC O6D4E8
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BCA2443;
OS Erwina carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
  Hoileva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
  Atkin R., Baason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
  Fraser A., Hance Z., Hauser H., Jagsels K., Moule S., Norczyczak H.,
  Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
  Salmon G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwina
  carotovora subsp. atroseptica and characterization of virulence
  factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
DR EMBL; BX950851; CAG75345.1; -; Genomic_DNA.
DR SMR; O6D4E8; 2-207.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 212 AA; 22892 MW; A67813005F32B150 CRC64;

Query Match 76.1%; Score 89; DB 2; Length 212;
Best Local Similarity 73.9%; Pred. No. 9.8e-06;
Matches 17; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCGSARLVHQLA 23

```

KW Complete proteome: Hypothetical protein.
SQ SEQUENCE 208 AA; 22730 MW; D0CC2B3BDE499723 CRC64;

Query Match 83.8%; Score 98; DB 2; Length 208;
Best Local Similarity 82.6%; Pred. No. 3.5e-07;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVQSARLVHOLA 23
Db 4 NYHDIYALAGVQSARLVHOLA 26

RESULT 5

YCFC_ECOL6
ID YCFC_ECOL6 STANDARD; PRT; 213 AA.

AC 08X736;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocNames=z1861, Ecol604;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=83334;

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RA STRAIN=0157:H7 / EDL933 / ATCC 700927 / EHEC;
RA MEDLINE=21156211; PubMed=11258796;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., D'Amalante E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RA Nature 409:529-533(2001).

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RA STRAIN=0157:H7 / Sakai / RIMD 0509952 / EHEC;
RA MEDLINE=21156211; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).

-1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; AE005174; AAG55958.1; -; Genomic DNA.
CC EMBL; BA000007; BAB35027.1; -; Genomic DNA.

DR PIR; B85687; B85687.
DR PIR; D90829; D90829.
DR SMR; Q8X736; 2-213.
DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome: Hypothetical protein.
SQ SEQUENCE 213 AA; 22947 MW; E26EP9676C3844E2 CRC64;

Query Match 83.8%; Score 98; DB 1; Length 213;
Best Local Similarity 82.6%; Pred. No. 3.5e-07;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVQSARLVHOLA 23

Db 4 NYHDIYALAGVQSARLVHOLA 26

RESULT 6

YCFC_ECOL6
ID YCFC_ECOL6 STANDARD; PRT; 213 AA.

AC 08F1B7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 47, Last annotation update)
DE Hypothetical UPF0274 protein ycfC.
GN Name=ycfC; OrderedLocNames=c1511;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_Taxid=217921;

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RA STRAIN=06:H1 / CFT073 / ATCC 700928 / UPEC;
RA MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.U., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RA "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

-1- SIMILARITY: Belongs to the UPF0274 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; AE016759; AAN79980.1; ALT_INIT; Genomic DNA.
CC SMR; Q8F1B7; 2-213.

DR HAMAP; MF_00695; -; 1.
DR InterPro; IPR007451; DUF489.
DR Pfam; PF04356; DUF489; 1.
KW Complete proteome: Hypothetical protein.
SQ SEQUENCE 213 AA; 22934 MW; 71C8P96F8BA37BB CRC64;

Query Match 83.8%; Score 98; DB 1; Length 213;
Best Local Similarity 82.6%; Pred. No. 3.5e-07;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVQSARLVHOLA 23
Db 4 NYHDIYALAGVQSARLVHOLA 26

RESULT 7

YCFC_ECOL1
ID YCFC_ECOL1 STANDARD; PRT; 213 AA.

AC P25746;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hypothetical UPF0274 protein ycfC (ORF-23).
GN Name=ycfC; OrderedLocNames=b1132;
OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.

OX NCBI_Taxid=562;

NUCLEOTIDE SEQUENCE.

RA STRAIN=K12;
RA Green S.M., Drabble W.T.;
RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
RP [2]

NUCLEOTIDE SEQUENCE.
RC STRAIN=K12;

RA MEDLINE=92104952; PubMed=1729205;
 RA He B., Smith J.M., Zaikin H.;
 RT "Escherichia coli purb gene: cloning, nucleotide sequence, and
 RT regulation by purR.";
 RL J. Bacteriol. 174:130-136(1992).
 RN (3)
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose J.D.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Naishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horichi T.;
 RT "A 718-Kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map.";
 RL DNA Res. 3:137-155(1996).
 CC -1 SIMILARITY: Belongs to the UPF0274 family.
 CC -1 CAUTION: Ref.2 sequence differs from that shown due to
 CC frameshifts.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC
 CC EMBL, X59307; CAA11995.1; -; Genomic DNA.
 DR EMBL, M74924; -; NOT ANNOTATED; CDS; Genomic DNA.
 DR EMBL, U00096; AAC74216.1; -; Genomic DNA.
 DR EMBL, D90748; BAA35954.1; -; Genomic DNA.
 DR EMBL, D90749; BAA35963.1; -; Genomic DNA.
 DR PIR, S19211; S19211.
 DR PDB, 1QZ4; X-ray; A=2-213.
 DR PDB, 1SD1; X-ray; A=2-213.
 DR Echobase; EBI121; -;
 DR Ecogene; EGI1345; ycfC.
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 DR 3D-structure; Complete proteome; Hypothetical protein.
 KW SEQUENCE 213 AA; 22948 MW; E26EF9698C3CEB42 CRC64;
 SQ

Query Match 83.8%; Score 98; DB 1; Length 213;
 Best Local Similarity 82.6%; Pred. No. 3.6e-07;
 Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NYHIVLALAGVCGSARLVHQLA 23
 DB 4 NYHIVLALAGVCGSARLVHQLA 26

RESULT 8
 YCFC SHIFL STANDARD; PRT; 213 AA.
 AC 0831F8;
 DT 10-OCT-2003 (Rel. 42, Last Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC.
 GN Name=ycfC; OrderedLocustNames=SF1151, S1234;
 OS Shigella flexneri.

CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Shigella.
 CC NCBI_TaxID=623;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkt566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qian B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RX DOI=10.1128/JMI.71.5.2775-2786.2003;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 CC -1 SIMILARITY: Belongs to the UPF0274 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC
 CC EMBL, AE005674; AAN42768.1; ALT INIT; Genomic DNA.
 DR EMBL, AB016982; AAP16657.1; -; Genomic DNA.
 DR SMR, Q83LF8; 2-213.
 DR HAMAP; MF_00695; -; 1.
 DR InterPro; IPR007451; DUF489.
 DR Pfam; PF04356; DUF489; 1.
 DR Complete proteome; Hypothetical protein.
 KW SEQUENCE 213 AA; 22894 MW; 0B0BEC69C29FBA4D CRC64;
 SQ

Query Match 83.8%; Score 98; DB 1; Length 213;
 Best Local Similarity 82.6%; Pred. No. 3.6e-07;
 Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NYHIVLALAGVCGSARLVHQLA 23
 DB 4 NYHIVLALAGVCGSARLVHQLA 26

RESULT 9
 YCFC SALTI STANDARD; PRT; 213 AA.
 AC 0827H0;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0274 protein ycfC.
 GN Name=ycfC; OrderedLocustNames=STY1273, t1687;
 OS Salmonella typhi.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Salmonella.
 CC NCBI_TaxID=601;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 RA Pakshill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 10:46:57 ; Search time 45 Seconds
(without alignments)

42.256 Million cell updates/sec

Title: US-10-501-838a-29

Sequence: 1 NYHDIYALAGVCGSARLVHQLA 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5.COMB.rep:*
2: /cgn2_6/prodata/1/1aa/6.COMB.rep:*
3: /cgn2_6/prodata/1/1aa/H.COMB.rep:*
4: /cgn2_6/prodata/1/1aa/PCTUS.COMB.rep:*
5: /cgn2_6/prodata/1/1aa/RE.COMB.rep:*
6: /cgn2_6/prodata/1/1aa/backfill.rep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	82.1	217	US-09-489-039A-7789	Sequence 7789, Ap
2	79	67.5	216	US-09-543-681A-7993	Sequence 7993, Ap
3	50	42.7	309	US-09-902-540-12828	Sequence 12828, A
4	49.5	42.3	367	US-09-404-296B-6	Sequence 6, Appl1
5	49	41.9	138	US-09-270-767-58510	Sequence 58510, A
6	49	41.9	138	US-09-270-767-43172	Sequence 43172, A
7	46	39.3	255	US-09-328-352-4347	Sequence 4347, Ap
8	44	37.6	338	US-09-543-681A-7046	Sequence 7046, Ap
9	43	36.8	483	US-09-252-991A-18354	Sequence 18354, A
10	43	36.8	625	US-09-252-991A-20865	Sequence 20865, A
11	42.5	36.3	158	US-08-828-832-3	Sequence 3, Appl1
12	42.5	36.3	158	US-09-486-632C-16	Sequence 16, Appl1
13	42.5	36.3	159	US-08-828-832-1	Sequence 1, Appl1
14	42.5	36.3	159	US-09-486-632C-15	Sequence 15, Appl1
15	42	35.9	246	US-09-199-637A-407	Sequence 407, App
16	41.5	35.5	240	US-09-194-905-8	Sequence 8, Appl1
17	41.5	35.5	240	US-09-922-683-8	Sequence 8, Appl1
18	41	35.0	205	US-09-902-540-15141	Sequence 15141, A
19	41	35.0	287	US-08-252-991A-29808	Sequence 29808, A
20	41	35.0	319	US-09-270-767-45362	Sequence 45362, A
21	41	35.0	436	US-09-605-703B-2864	Sequence 2864, Ap
22	41	35.0	470	US-09-769-863-20	Sequence 20, Appl1
23	41	35.0	624	US-09-252-991A-25383	Sequence 25383, A
24	41	35.0	1037	US-09-902-540-9845	Sequence 9845, Ap
25	40.5	34.6	375	US-09-498-520A-24	Sequence 24, Appl1
26	40.5	34.6	551	US-09-252-991A-20014	Sequence 20014, A
27	40.5	34.6	762	US-09-228-986-114	Sequence 114, App

28	40.5	34.6	762	2	US-10-101-466A-114	Sequence 114, App
29	40	34.2	119	2	US-10-104-047-2840	Sequence 2840, Ap
30	40	34.2	162	2	US-09-540-236-3655	Sequence 3655, Ap
31	40	34.2	183	2	US-09-134-000C-4230	Sequence 4230, Ap
32	40	34.2	231	2	US-09-830-433A-32	Sequence 32, Appl1
33	40	34.2	303	2	US-09-543-681A-7924	Sequence 7924, Ap
34	40	34.2	306	2	US-10-686-944-5	Sequence 5, Appl1
35	40	34.2	342	2	US-09-252-991A-27215	Sequence 27215, A
36	40	34.2	369	1	US-07-937-609-19	Sequence 19, Appl1
37	40	34.2	369	2	US-08-029-170-19	Sequence 19, Appl1
38	40	34.2	369	2	US-09-443-745-19	Sequence 19, Appl1
39	40	34.2	388	2	US-09-940-921B-7	Sequence 7, Appl1
40	40	34.2	398	2	US-09-940-921B-9	Sequence 9, Appl1
41	40	34.2	420	1	US-08-846-762-73	Sequence 73, Appl1
42	40	34.2	462	1	US-08-471-033-20	Sequence 20, Appl1
43	40	34.2	462	1	US-08-471-044-20	Sequence 20, Appl1
44	40	34.2	462	1	US-08-463-483A-20	Sequence 20, Appl1
45	40	34.2	462	1	US-08-471-046A-20	Sequence 20, Appl1

ALIGNMENTS

RESULT 1
US-09-489-039A-7789
Sequence 7789, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 7789
LENGTH: 217
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7789

Query Match 82.1%; Score 96; DB 2; Length 217;
Best Local Similarity 82.6%; Pred. No. 1e-07;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

CY 1 NYHDIYALAGVCGSARLVHQLA 23
DB 8 NYHDIYALAGVCGSARLVHQLA 30

RESULT 2
US-09-543-681A-7993
Sequence 7993, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7993
LENGTH: 216
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7993

Query Match 67.5%; Score 79; DB 2; Length 216;

Db 41 LALAGVFOATQUTMNTA 57

RESULT 8

US-09-543-681A-7046
Sequence 7046; Application US/09543681A

Patent No. 6605709

GENERAL INFORMATION:

APPLICANT: GARY BRETON

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.1002-001

CURRENT APPLICATION NUMBER: US/09/543,681A

CURRENT FILING DATE: 2000-04-05

PRIOR APPLICATION NUMBER: US 60/128,706

PRIOR FILING DATE: 1999-04-09

NUMBER OF SEQ ID NOS: 8344

SEQ ID NO 7046

LENGTH: 398

TYPE: PRT

ORGANISM: Proteus mirabilis

US-09-543-681A-7046

Query Match 37.6%; Score 44; DB 2; Length 398;

Best Local Similarity 34.5%; Pred. No. 53;

Matches 10; Conservative 4; Mismatches 5; Indels 10; Gaps 1;

QY 3 HDIVLALAG-----VCOSARLVHQ 21

Db 253 HDAAVLRLGKIKTSLIMEXICOMALIAQ 281

RESULT 9

US-09-252-991A-18354

Sequence 18354; Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfeld et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 18354

LENGTH: 483

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-18354

Query Match 36.8%; Score 43; DB 2; Length 483;

Best Local Similarity 33.3%; Pred. No. 96;

Matches 6; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 5 IVLALAGVCOSARLVHQ 22

Db 348 IISASGMCDCAGRIIRHL 365

RESULT 10

US-09-252-991A-20865

Sequence 20865; Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfeld et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 20865

LENGTH: 625

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-20865

Query Match 36.8%; Score 43; DB 2; Length 625;

Best Local Similarity 37.5%; Pred. No. 138+02;

Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 5 IVLALAGVCOSARLVH 20

Db 498 VIIASGMCCTGRIYH 513

RESULT 11

US-08-828-832-3

Sequence 3; Application US/08828832

Patent No. 5827711

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Shah, Purni

TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSER: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/828,832

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0250 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 158 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCES:

LIBRARY: GenBank

CLONE: 1575011

US-08-828-832-3

Query Match 36.3%; Score 42.5; DB 1; Length 158;

Best Local Similarity 31.8%; Pred. No. 33;

Matches 7; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

QY 1 NYHDIVLALAGVCOSARLVHQ 22

```
Db      142 NYHDV-----GICKAVAMLMKTL 158

RESULT 12
US-09-496-632C-16
; Sequence 16, Application US/09496632C
; Patent No. 6468789
; GENERAL INFORMATION:
; APPLICANT: BAYSAL, Bora E.
; APPLICANT: FERRELL, Robert E.
; APPLICANT: DEVLIN, Bernie J.
; APPLICANT: WILLETT-BROZICK, Joan E.
; TITLE OF INVENTION: OXYGEN SENSING AND HYPOXIC SELECTION FOR TUMORS
; FILE REFERENCE: 99-484-US
; CURRENT APPLICATION NUMBER: US/09/496,632C
; CURRENT FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-496-632C-16

Query Match      36.3%; Score 42.5; DB 2; Length 158;
Best Local Similarity 31.8%; Pred. No. 33;
Matches 7; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

Cy      1 NYHDIVLAGVCOSARLVHQL 22
      ||||:|:|:|:|:|
      142 NYHDV-----GICKAVAMLMKTL 158

RESULT 13
US-08-828-832-1
; Sequence 1, Application US/08828832
; Patent No. 5827711
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Shah, Putvi
; TITLE OF INVENTION: NOVEL SUCCINATE DEHYDROGENASE SUBUNIT
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,832
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0250 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 159 amino acids

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: 2454416
US-08-828-832-1

Query Match      36.3%; Score 42.5; DB 1; Length 159;
Best Local Similarity 31.8%; Pred. No. 33;
Matches 7; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

Cy      1 NYHDIVLAGVCOSARLVHQL 22
      ||||:|:|:|:|:|
      143 NYHDV-----GICKAVAMLMKTL 159

RESULT 14
US-09-496-632C-15
; Sequence 15, Application US/09496632C
; Patent No. 6468789
; GENERAL INFORMATION:
; APPLICANT: BAYSAL, Bora E.
; APPLICANT: FERRELL, Robert E.
; APPLICANT: DEVLIN, Bernie J.
; APPLICANT: WILLETT-BROZICK, Joan E.
; TITLE OF INVENTION: OXYGEN SENSING AND HYPOXIC SELECTION FOR TUMORS
; FILE REFERENCE: 99-484-US
; CURRENT APPLICATION NUMBER: US/09/496,632C
; CURRENT FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-496-632C-15

Query Match      36.3%; Score 42.5; DB 2; Length 159;
Best Local Similarity 31.8%; Pred. No. 33;
Matches 7; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

Cy      1 NYHDIVLAGVCOSARLVHQL 22
      ||||:|:|:|:|:|
      143 NYHDV-----GICKAVAMLMKTL 159

RESULT 15
US-09-199-637A-407
; Sequence 407, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT APPLICATION NUMBER: US/09/199,637A
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 407
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
```

US-09-199-637A-407

Query Match 35.9%; Score 42; DB 2; Length 246;
Best Local Similarity 42.9%; Pred. No. 66;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 YHDIVLALAGVQGS 15
:|::|:|:
Db 41 HHDVLAGPAGICLS 54

Search completed: January 23, 2006, 11:33:40
Job time : 45 secs

This Page Blank (over)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Comphen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 12:09:45 ; Search time 9 seconds
(without alignments)
25.898 Million cell updates/sec

Title: US-10-501-838a-29

Perfect score: 117
Sequence: 1 NYHDIYALAGVCSARLVHQLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 70606 seqs, 10133881 residues

Total number of hits satisfying chosen parameters: 70606

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/pubppaa/US09_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubppaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubppaa/US05_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubppaa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	35.9	42	6 US-10-512-295A-3	Sequence 3, Appl1
2	40	34.2	306	7 US-11-017-550-45	Sequence 45, Appl1
3	40	34.2	306	7 US-11-014-402-5	Sequence 5, Appl1
4	40	34.2	306	7 US-11-137-850-6	Sequence 6, Appl1
5	40	34.2	388	7 US-11-046-668-7	Sequence 7, Appl1
6	40	34.2	388	7 US-11-046-668-9	Sequence 9, Appl1
7	39.5	33.8	365	7 US-11-000-463-241	Sequence 241, App
8	39.5	33.8	365	7 US-11-000-463-713	Sequence 713, App
9	38.5	32.9	305	6 US-10-520-820-1	Sequence 1, Appl1
10	38	32.5	185	6 US-10-453-372-896	Sequence 896, App
11	38	32.5	184	6 US-10-453-372-894	Sequence 894, App
12	38	32.5	218	6 US-10-453-372-892	Sequence 892, App
13	37	31.6	228	6 US-10-467-657-1242	Sequence 1242, App
14	37	31.6	307	6 US-10-793-626-684	Sequence 684, App
15	37	31.6	412	6 US-10-979-821-8	Sequence 8, Appl1
16	37	31.6	412	6 US-11-114-922-8	Sequence 8, Appl1
17	37	31.6	2644	6 US-10-770-726-45	Sequence 45, Appl1
18	36.5	31.2	447	6 US-10-858-730-219	Sequence 219, App
19	36.5	31.2	465	7 US-11-082-389-284	Sequence 284, App
20	36	30.8	204	6 US-10-858-730-95	Sequence 95, Appl1
21	36	30.8	231	7 US-11-082-389-242	Sequence 242, App
22	36	30.8	322	6 US-10-689-742-46	Sequence 46, Appl1
23	36	30.8	379	6 US-10-858-730-24	Sequence 24, App
24	36	30.8	379	6 US-10-858-730-281	Sequence 281, App
25	36	30.8	379	6 US-10-858-730-285	Sequence 285, App

25	36	30.8	379	6	US-10-858-730-290	Sequence 290, App
27	36	30.8	502	6	US-10-063-703-158	Sequence 158, App
28	36	30.8	502	7	US-11-102-240-158	Sequence 158, App
29	36	30.8	513	7	US-11-055-822-348	Sequence 348, App
30	36	30.8	810	6	US-10-453-372-1116	Sequence 1116, App
31	35.5	30.3	162	6	US-10-467-657-2510	Sequence 2510, App
32	35.5	30.3	327	7	US-11-024-959-455	Sequence 455, App
33	35.5	30.3	652	6	US-10-873-528-26	Sequence 26, Appl1
34	35	29.9	113	6	US-10-467-657-2448	Sequence 2448, App
35	35	29.9	180	5	US-09-978-360A-463	Sequence 463, App
36	35	29.9	250	7	US-11-051-670-2	Sequence 2, Appl1
37	35	29.9	312	7	US-11-156-084-236	Sequence 236, App
38	35	29.9	312	7	US-11-156-084-308	Sequence 308, App
39	35	29.9	319	6	US-10-467-657-3252	Sequence 3252, App
40	35	29.9	346	6	US-11-186-284-189	Sequence 189, App
41	35	29.9	433	6	US-10-131-826A-334	Sequence 334, App
42	35	29.9	519	6	US-10-523-477-8	Sequence 8, Appl1
43	35	29.9	2556	7	US-11-050-346-67	Sequence 67, Appl1
44	35	29.9	6893	7	US-11-205-109-14	Sequence 14, Appl1
45	34.5	29.5	447	7	US-11-055-822-94	Sequence 94, Appl1

ALIGNMENTS

```

RESULT 1
US-10-512-295A-3
; Sequence 3, Application US/10512295A
; Publication No. US20050245727A1
GENERAL INFORMATION:
; APPLICANT: Balzer, Lara
; APPLICANT: Dolphin, Gunnar
; APPLICANT: Liedberg, Bo
; APPLICANT: Lundstrom, Ingemar
; TITLE OF INVENTION: NOVEL POLYPEPTIDE SCAFFOLDS AND USE THEREOF
; FILE REFERENCE: 5848.18USWO
; CURRENT APPLICATION NUMBER: US/10/512,295A
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/SE03/00507
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: SE 0200968-6
; PRIOR FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide KES3
US-10-512-295A-3
Query Match          35.9%; Score 42; DB 6; Length 42;
Best Local Similarity 52.9%; Pred. No. 1.3;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      7 LALAGVCSARLVHQLA 23
      |||||:|:|
Db      16 LAARGCDAAQLAEQLA 32

RESULT 2
US-11-017-550-45
; Sequence 45, Application US/11017550
; Publication No. US20050250183A1
GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Schultz, Peter G
; APPLICANT: Wang, Lei
; APPLICANT: Anderson, John C
; APPLICANT: Chiu, Jason
; APPLICANT: Liu, David R
; APPLICANT: Meglery, Thomas

```

```

APPLICANT: IRM, LLC
APPLICANT: Xie, Jiaming
APPLICANT: Wang, Lei
APPLICANT: Wu, Ning
APPLICANT: Schultz, Peter G
APPLICANT: Sprengon, Glen
TITLE OF INVENTION: SITE SPECIFIC INCORPORATION OF HEAVY ATOM-CONTAINING UNNATURAL
FILE REFERENCE: 54-000920US
CURRENT FILING DATE: 2005-05-24
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.2
SEQ ID NO 6
LENGTH: 306
TYPE: PRT
ORGANISM: Artificial
FEATURES:
OTHER INFORMATION: mutant synthetase
US-11-137-850-6

Query Match          34.2%; Score 40; DB 7; Length 306;
Best local Similarity 39.1%; Pred. No. 22;
Matches 9; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

CY      1 NYHDIYALAGVCQARLVHOLA 23
      :||:|:|:|:|:|:|:|:|:|:|
Db      160 HYHGVDAVAGGMEQ-RKIHMLA 180

RESULT 5
US-11-046-668-7
Sequence 7, Application US/11046668
Publication No. US20050244855A1
GENERAL INFORMATION:
APPLICANT: Friddele, Carl Johan
APPLICANT: Hilbun, Erin
APPLICANT: Nepomnichy, Boris
APPLICANT: Hu, Yi
TITLE OF INVENTION: Novel Human Kinase Proteins and Polynucleotides Encoding the Same
FILE REFERENCE: Lex-0227-USA
CURRENT APPLICATION NUMBER: US/11/046,668
CURRENT FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: US/09/940,921
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/229,280
PRIOR FILING DATE: 2000-08-31
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 388
TYPE: PRT
ORGANISM: homo sapiens
US-11-046-668-7

Query Match          34.2%; Score 40; DB 7; Length 388;
Best local Similarity 31.6%; Pred. No. 29;
Matches 6; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

CY      4 DIVTALAGVCQARLVHOL 22
      |:|:|:|:|:|:|:|:|:|:|
Db      203 DTLFMKQICEGIRHMHOM 221

RESULT 6
US-11-046-668-9
Sequence 9, Application US/11046668
Publication No. US20050244855A1
GENERAL INFORMATION:
APPLICANT: Friddele, Carl Johan
APPLICANT: Hilbun, Erin
APPLICANT: Nepomnichy, Boris
APPLICANT: Hu, Yi

```



```

; LENGTH: 305
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-520-820-1

Query Match      32.9% Score 38.5; DB 6; Length 305;
Best Local Similarity 47.6%; Pred. No. 39;
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Qy      1 NYHDIYALAGVCSARLVHOLA 23
      ||| : ||| : ||| : |||
Db      68 NY-DAVIDAAGLVKSALVTRLA 89

RESULT 10
US-10-453-372-896
; Sequence 896, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
; APPLICANT: Alsebrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Cuiaseqblast version 0.1
; SEQ ID NO 896
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-896

Query Match      32.5% Score 38; DB 6; Length 185;
Best Local Similarity 47.6%; Pred. No. 28;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy      3 HDIVIALAGVCSARLVHOLA 23
      : ||| : ||| : ||| : |||
Db      67 YDSLALDGHIGSARALMVVA 87

RESULT 11
US-10-453-372-894
; Sequence 894, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
; APPLICANT: Alsebrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
```

```

; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Cuiaseqblast version 0.1
; SEQ ID NO 892
; LENGTH: 218
; TYPE: PRT

Qy      3 HDIVIALAGVCSARLVHOLA 23
      : ||| : ||| : ||| : |||
Db      47 YDSLALDGHIGSARALMVVA 67

RESULT 12
US-10-453-372-892
; Sequence 892, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:
; APPLICANT: Alsebrook, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-589 A
; CURRENT APPLICATION NUMBER: US/10/453,372
; PRIOR FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: 09/789390
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/823187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195792
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 09/839446
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/199476
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: 09/863776
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/208263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/939398
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: Cuiaseqblast version 0.1
; SEQ ID NO 892
; LENGTH: 218
; TYPE: PRT
```

ORGANISM: Homo sapiens
US-10-453-372-892

Query Match 32.5%; Score 38; DB 6; Length 218;
Best Local Similarity 47.6%; Pred. No. 33;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 HDIVLAGVCOSARLVHQLA 23
DB 71 YDSLALDGHIGSARALMVVA 91

RESULT 13
US-10-467-657-1242
Sequence 1242, Application US/10467657
Publication No. US20050260581A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: FONTANA Maria Rita
APPLICANT: PIZZA Mariagrazia
APPLICANT: MASIGNANI Vega
APPLICANT: MONACI Elisabetta
TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/467,657
CURRENT FILING DATE: 2003-08-11
PRIOR APPLICATION NUMBER: GB-0103424.8
PRIOR FILING DATE: 2001-02-12
NUMBER OF SEQ ID NOS: 9218
SOFTWARE: SeqWin99, version 1.04
SEQ ID NO 1242
LENGTH: 228
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1242

Query Match 31.6%; Score 37; DB 6; Length 228;
Best Local Similarity 47.1%; Pred. No. 50;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 IYVLAGVCOSARLVHQ 21
DB 105 LVLSVSGLFYFARLVQ 121

RESULT 14
US-10-793-626-684
Sequence 684, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 684
LENGTH: 307
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-10-793-626-684

Query Match 31.6%; Score 37; DB 6; Length 307;
Best Local Similarity 30.3%; Pred. No. 68;
Matches 10; Conservative 5; Mismatches 8; Indels 10; Gaps 1;
QY 1 NYHDIVLAGV-----COSARLVHQLA 23

DB 54 NHKQVTEAMKGIIDIAVYLDPTKSAKLTHTA 86

RESULT 15
US-10-979-821-8
Sequence 8, Application US/10979821
Publication No. US20050244937A1
GENERAL INFORMATION:
APPLICANT: ABRAHAM, TIMOTHY W.
APPLICANT: CAMERON, DOUGLAS C.
APPLICANT: HICKS, PAULA M.
APPLICANT: MCPARLAN, SARA C.
APPLICANT: MILLIS, JIM
APPLICANT: ROSAZZA, JACK
APPLICANT: ZHAO, LISHAN
APPLICANT: WEINER, DAVID P.
TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
TITLE OF INVENTION: PRODUCTION OF MONATIN AND ITS PRECURSORS
FILE REFERENCE: 023829-0390
CURRENT APPLICATION NUMBER: US/10/979,821
CURRENT FILING DATE: 2004-11-03
PRIOR APPLICATION NUMBER: 10/422,366
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 60/374,831
PRIOR FILING DATE: 2002-04-23
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 8
LENGTH: 412
TYPE: PRT
ORGANISM: Leishmania major
US-10-979-821-8

Query Match 31.6%; Score 37; DB 6; Length 412;
Best Local Similarity 33.3%; Pred. No. 93;
Matches 7; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 HDIVLAGVCOSARLVHQLA 23
DB 375 HNIFIVSGRANAGLTHETA 395

Search completed: January 23, 2006, 12:13:23
Job time : 9 secs

This Page Blank (except)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2006, 10:28:43 ; Search time 134 Seconds
(without alignments)
75.416 Million cell updates/sec

Title: US-10-501-838a-29

Sequence: 1 NYHDIYALAGVCGSARLVHOLA 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*

1: genebep1980a:*
2: genebep1990a:*
3: genebep2000a:*
4: genebep2001a:*
5: genebep2002a:*
6: genebep2003a:*
7: genebep2004a:*
8: genebep2005a:*
9: genebep2006a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	100.0	23	6	ADBI6912 Haemophil
2	117	100.0	23	9	AEBO8246 Haemophil
3	117	100.0	29	6	ADBI6916 Escherich
4	117	100.0	29	9	AEBO8250 Penetrati
5	114	97.4	23	6	ADBI6886 Haemophil
6	114	97.4	23	9	AEBO8218 Haemophil
7	114	97.4	205	4	AAW50230 Haemophil
8	114	97.4	205	4	AEBO8261 Haemophil
9	98	83.8	23	6	ADBI6888 Escherich
10	98	83.8	23	9	AEBO8220 Escherich
11	98	83.8	24	6	ADBI6908 Escherich
12	98	83.8	24	9	AEBO8242 Escherich
13	98	83.8	25	6	ADBI6922 Escherich
14	98	83.8	25	6	ADBI6906 Escherich
15	98	83.8	26	6	ADBI6904 E coli pe
16	98	83.8	29	6	ADBI6921 E coli pe
17	98	83.8	30	6	ADBI6903 E coli pe
18	98	83.8	30	6	ADBI6905 Escherich
19	98	83.8	30	9	AEBO8239 Penetrati
20	98	83.8	31	6	ADBI6918 Escherich
21	98	83.8	31	9	AEBO8252 Penetrati
22	98	83.8	213	9	AEBO8263 Escherich
23	96	82.1	217	7	ABO61272 Klebsiell
24	92	78.6	23	6	ADBI6909 Escherich

25	92	78.6	23	9	AEBO8243 Escherich
26	92	78.6	30	6	ADBI6919 Escherich
27	92	78.6	30	9	AEBO8253 Penetrati
28	89	76.1	23	6	ADBI6887 Pasteurel
29	89	76.1	23	9	AEBO8219 Pasteurel
30	89	76.1	203	9	AEBO8262 Pasteurel
31	85	72.6	22	6	ADBI6910 Escherich
32	85	72.6	22	9	AEBO8244 Escherich
33	85	72.6	29	6	ADBI6820 Escherich
34	85	72.6	29	9	AEBO8254 Escherich
35	81	69.2	209	6	ABM67069 Buchnera
36	79	67.5	216	7	ADFO7708 Buchnera
37	73	62.4	23	6	ADBI6890 Buchnera
38	73	62.4	23	9	AEBO8222 Buchnera
39	73	62.4	211	9	AEBO8265 Buchnera
40	56	47.9	23	6	ADBI6889 Vibrio ch
41	56	47.9	23	9	AEBO8221 Vibrio ch
42	56	47.9	204	9	AEBO8264 Vibrio ch
43	50	42.7	309	9	ABM93629 M. xanthu
44	49	41.9	468	4	ABR61357 Drosophil
45	48	41.0	23	6	ADBI6892 Xylella f

ALIGNMENTS

RESULT 1
ADBI6912 standard; peptide; 23 AA.

ADBI6912;

20-NOV-2003 (first entry)

Haemophilus influenzae ORF HI0638 penetrating peptide 34.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;

infertility; hormone; vitamin deficiency; neurodegenerative;

cardiovascular; haematological; endocrine disorder; obesity;

neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;

osteopathic; cytoskeletal; nootropic.

Haemophilus influenzae.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-1B000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating

or preventing e.g. endocrine disorders.

Claim 2; Page 15; 60pp; English.

This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony

CC stimulating factor (GM-CSF), enkephalin, dargirin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cyostatic or neurotropic activities. This peptide is from ORF HI0638 of
 CC Haemophilus influenzae and is penetrating peptide 34 of the invention.
 CC
 CC Sequence 23 AA;
 CC
 CC Query Match 100.0%; Score 117; DB 6; Length 23;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e-11;
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 NYHDIIVLALAGVCSARLVHQLA 23
 CC 1 NYHDIIVLALAGVCSARLVHQLA 23
 CC
 CC RESULT 2
 CC AEB08246
 CC ID AEB08246 standard; peptide; 23 AA.
 CC
 CC AC AEB08246;
 CC
 CC DT 25-AUG-2005 (first entry)
 CC
 CC DE Haemophilus influenzae ORF HI0638 penetrating peptide 34, SEQ ID NO: 29.
 CC
 CC KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degenerative; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntingtons chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; anti anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virocidic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cyostatic;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 CC
 CC KW Haemophilus influenzae.
 CC OS US2005136103-A1.
 CC
 CC PN 23-JUN-2005.
 CC
 CC PD 16-SEP-2004; 2004US-00942300.
 CC
 CC PP 17-SEP-2003; 2003US-00664989.
 CC PR 17-SEP-2003; 2003US-00665184.
 CC PR 17-SEP-2003; 2003US-0503615P.
 CC
 CC PA (BENS/) BEN-SASSON S A.
 CC (COHEN/) COHEN E.
 CC
 CC PI Ben-Sasson SA, Cohen E;
 CC
 CC DR WPI; 2005-444089/45.
 CC
 CC XX Composition used for translocating effectors across barrier such as
 PT epithelial cells during treatment of e.g. endocrine disorders comprises

PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 46; SEQ ID NO 29; 59pp; English.
 PS
 XX The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transcellular delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Haemophilus influenzae ORF
 CC HI0638 penetrating peptide. This sequence is used in the effective
 CC translocation of aminoglycoside antibiotics and antifungal agents across
 CC an epithelial barrier.
 CC
 CC SQ Sequence 23 AA;
 CC
 CC Query Match 100.0%; Score 117; DB 9; Length 23;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e-11;
 CC Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 NYHDIIVLALAGVCSARLVHQLA 23
 CC 1 NYHDIIVLALAGVCSARLVHQLA 23
 CC
 CC QY
 CC DB 1 NYHDIIVLALAGVCSARLVHQLA 23
 CC
 CC RESULT 3
 CC ADB16916
 CC ID ADB16916 standard; peptide; 29 AA.
 CC
 CC AC ADB16916;
 CC
 CC DT 20-NOV-2003 (first entry)
 CC
 CC DE Escherichia coli peptide 3 coupled to imaging compound linker, IBM-006.
 CC
 CC KW epithelial; endothelial; tight junction; diabetes; infertility; hormone;
 KW vitamin deficiency; neurodegenerative; cardiovascular; haematological;
 KW cardiac; antiarteriosclerotic; osteopathic; cyostatic; neurotropic;
 KW imaging linker; penetrating peptide; IBM-006.
 CC
 CC KW Synthetic.
 CC OS Escherichia coli.
 CC
 CC OS Key 1 Location/Qualifiers
 CC FH Modified-site 1 /note= "N-terminal acetyl"
 CC FT Modified-site 29 /note= "C-terminal amide"
 CC FT
 CC PN WO2003066859-A2.
 CC
 CC PD 14-AUG-2003.
 CC
 CC PP 07-FEB-2003; 2003WO-IB000968.
 CC
 CC PR 07-FEB-2002; 2002US-0355396P.
 CC
 CC PR

PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

XX Ben-Sasson SA, Cohen E;

XX WPI; 2003-697452/66.

PT New penetrating peptide, useful for preparing a composition for treating

PT or preventing e.g. endocrine disorders.

PS Example 3; Page 40; 60pp; English.

CC This invention relates to a novel peptide sequences capable of
CC translocating across a biological barrier. Furthermore, it refers to
CC methods that use these peptides to facilitate penetration of a
CC biologically active effector molecule such as a drug or other therapeutic
CC agent across biological barriers e.g. epithelial or endothelial cells
CC sealed by tight junctions. This peptide is derived from a bacterial
CC toxin, an integral membrane or extracellular protein and can comprise an
CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
CC or enzyme. The effector molecule, however, can comprise for example
CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
CC factors. The penetrating peptide is useful for the treatment of various
CC conditions including diabetes, infertility, hormone and vitamin and
CC deficiencies, neurodegenerative, cardiovascular, haematological and
CC endocrine disorders, as well as obesity and neoplastic disease.
CC Accordingly, the peptides of this invention can be used in compositions
CC that have neuroprotective, cardiant, antihypertensive, osteopathic,
CC cytoskeletal or neurotrophic activities. This peptide sequence is IBM-006,
CC which consists of the Escherichia coli penetrating peptide 3 coupled to
CC the imaging linker peptide used in an exemplification of the invention.

SEQ Sequence 29 AA;

Query Match 100.0%; Score 117; DB 6; Length 29;

Best Local Similarity 100.0%; Pred. No. 2.3e-11; Mismatches 0; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NYHDIYALAGVCSARLVHQLA 23

1 NYHDIYALAGVCSARLVHQLA 23

RESULT 4
AEB08250 standard; peptide; 29 AA.

AC AEB08250;

DT 25-AUG-2005 (first entry)

DE Penetrating peptide SEQ: 33 used in composition for mucosal vaccination.

XX Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;

XX antidiabetic; endocrine disease; gastrointestinal disease;

XX metabolic disorder; hormone deficiency; osteoporosis; ophthalmic;

XX degenerative; musculoskeletal disease; ocular disease; ophthalmological;

XX neurodegenerative disease; neuroprotective; Alzheimer's disease;

XX neurologic; neurological disease; parkinson's disease; antiparkinsonian;

XX dementia; multiple sclerosis; immune disorder; Huntington's chorea;

XX cardiovascular; genetic disorder; cardiovascular disease;

XX cardiovascular-gen.; atherosclerosis; antiatherosclerotic;

XX coronary artery disease; cardiant; vasotrophic; obesity; anorectic;

XX nutritional disorder; vitamin deficiency; renal disease; nephrotoxic;

XX genitourinary disease; hematological disease; antianemic; anemia;

XX autoimmune disease; immunosuppressive; immune deficiency;

XX immunostimulant; infectious disease; antimicrobial; infection;

XX erectile dysfunction; andrology; major depressive disorder;

XX antidepressant; psychiatric disorder; pain; analgesic;

XX bacterial infection; antibacterial; viral infection; virostatic;

XX fungal infection; fungicide; parasitic infection; antiparasitic;

XX renal failure; antiferility; antirheumatic; cytoskeletal;

XX antiinflammatory; hepatotropic; hepatitis B virus infection.

XX Unidentified.

XX Key Location/Qualifiers

XX Misc-difference 1 /note= "N-terminal acylated"

XX Misc-difference 29 /note= "Optionally C-terminal amide, optionally the free

XX amino group of lysine is acylated with a fatty acid"

XX US2005136103-A1.

XX 23-JUN-2005.

XX 16-SEP-2004; 2004US-00942300.

XX 17-SEP-2003; 2003US-00664989.

XX 17-SEP-2003; 2003US-00665184.

XX 17-SEP-2003; 2003US-0503615P.

XX (BENS/) BEN-SASSON S A.

XX (COHE/) COHEN E.

XX Ben-Sasson SA, Cohen E;

XX WPI; 2005-444089/45.

XX Claim 63; SEQ ID NO 33; 59pp; English.

CC The present invention relates to a pharmaceutical composition of
CC penetrating peptides for trans epithelial delivery of effector. The
CC invention comprises the effector sequentially coupled with a counter ion
CC and at least one hydrophobic agent, where the effector is selectively
CC encapsulated into a complex. The invention is useful for translocating
CC effectors across a biological barrier such as epithelial cells and
CC endothelial cells during treatment and prevention of disease or
CC pathological conditions (including endocrine disorders, diabetes,
CC infertility, hormone deficiencies, osteoporosis, ophthalmological
CC disorders, neurodegenerative disorders, Alzheimer's disease,
CC Parkinson's disease, multiple sclerosis, Huntington's disease,
CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
CC coagulable states, coronary disease, cerebrovascular events, metabolic
CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
CC hematological disorders, anemia of different entities, immunologic and
CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
CC infectious diseases, viral infections, bacterial infections, fungal
CC infections, parasitic infections, neoplastic diseases, multi-factorial
CC disorders, impotence, chronic pain, depression, different fibrosis states
CC and short stature) and for mucosal vaccination against anthrax and
CC hepatitis B. The present sequence is a penetrating peptide (IBM-006) used
CC in the composition for mucosal vaccination using a counter anion and a
CC penetrating peptide.

SEQ Sequence 29 AA;

Query Match 100.0%; Score 117; DB 9; Length 29;

Best Local Similarity 100.0%; Pred. No. 2.3e-11; Mismatches 0; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 NYHDIYALAGVCSARLVHQLA 23

1 NYHDIYALAGVCSARLVHQLA 23

RESULT 5

ADBI6886 standard; peptide; 23 AA.

XX ADBI6886;

XX 20-NOV-2003 (first entry)
 XX Haemophilus influenzae ORF HI0638 penetrating peptide 1.
 DE
 XX penetrating peptide; epithelial; endothelial; tight junction; diabetes;
 XX infertility; hormone; vitamin deficiency; neurodegenerative;
 XX cardiovascular; haematological; endocrine disorder; obesity;
 XX neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 XX osteopathic; cytoskeletal; noctropic.
 XX Haemophilus influenzae.
 OS
 XX WO2003066859-A2.
 XX
 XX 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-1B000968.
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2003-697452/66.
 XX
 XX New penetrating peptide, useful for preparing a composition for treating
 XX or preventing e.g. endocrine disorders.
 XX
 XX Claim 2; Page 14; 60pp; English.
 XX
 CC This invention relates to a novel peptide sequences capable of
 CC translocating across a biological barrier. Furthermore, it refers to
 CC methods that use these peptides to facilitate penetration of a
 CC biologically active effector molecule such as a drug or other therapeutic
 CC agent across biological barriers e.g. epithelial or endothelial cells
 CC sealed by tight junctions. This peptide is derived from a bacterial
 CC toxin, an integral membrane or extracellular protein and can comprise an
 CC anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 CC or enzyme. The effector molecule, however, can comprise for example
 CC insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
 CC stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 CC factors. The penetrating peptide is useful for the treatment of various
 CC conditions including diabetes, infertility, hormone and vitamin
 CC deficiencies, neurodegenerative, cardiovascular, haematological and
 CC endocrine disorders, as well as obesity and neoplastic disease.
 CC Accordingly, the peptides of this invention can be used in compositions
 CC that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 CC cytoskeletal or noctropic activities. This peptide is from ORF HI0638 of
 CC Haemophilus influenzae and is penetrating peptide 1 of the invention.
 XX
 XX Sequence 23 AA;
 SQ
 Query Match 97.4%; Score 114; DB 6; Length 23;
 Best Local Similarity 95.7%; Pred. No. 5.3e-11;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYHDIIVLALAGVCGSARLVHQLA 23
 DB 1 NYHDIIVLALAGVCGSARLVHQLA 23

RESULT 6
 AEB08218
 ID AEB08218 standard; peptide; 23 AA.
 XX
 XX AEB08218;
 AC
 XX
 XX 25-AUG-2005 (first entry)
 DT
 XX
 DE Haemophilus influenzae ORF HI0638 penetrating peptide 1, SEQ ID NO: 1.

KW Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
 KW antidiabetic; endocrine disease; gastrointestinal disease;
 KW metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
 KW degeneration; musculoskeletal disease; ocular disease; ophthalmological;
 KW neurodegenerative disease; neuroprotective; Alzheimer's disease;
 KW noctropic; neurological disease; parkinsons disease; antiparkinsonian;
 KW dementia; multiple sclerosis; immune disorder; Huntington's chorea;
 KW anticonvulsant; genetic disorder; cardiovascular disease;
 KW cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
 KW coronary artery disease; cardiac; vasotropic; obesity; anorectic;
 KW nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
 KW genitourinary disease; hematological disease; antianemic; anemia;
 KW autoimmune disease; immunosuppressive; immune deficiency;
 KW immunostimulant; infectious disease; antimicrobial; infection;
 KW erectile dysfunction; andrology; major depressive disorder;
 KW antidepressant; psychiatric disorder; pain; analgesic;
 KW bacterial infection; antibacterial; viral infection; virolytic;
 KW fungal infection; fungicide; parasitic infection; antiparasitic;
 KW renal failure; antifertility; antirheumatic; cytoskeletal;
 KW antiinflammatory; hepatotropic; hepatitis B virus infection.
 KW
 XX Haemophilus influenzae.
 OS
 XX US2005136103-A1.
 XX
 XX 23-JUN-2005.
 XX
 XX 16-SEP-2004; 2004US-00942300.
 XX
 XX 17-SEP-2003; 2003US-00664989.
 XX 17-SEP-2003; 2003US-00665184.
 XX 17-SEP-2003; 2003US-0503615P.
 XX
 XX (BENS/) BEN-SASSON S A.
 XX (COHE/) COHEN E.
 XX Ben-Sasson SA, Cohen E;
 XX WPI; 2005-444089/45.
 XX
 DR Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprises
 XX effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 46; SEQ ID NO 1; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transepithelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Haemophilus influenzae ORF
 CC HI0638 penetrating peptide. This sequence is used in the effective
 CC translocation of aminoglycoside antibiotics and antifungal agents across
 CC an epithelial barrier.
 XX
 XX Sequence 23 AA;

Query Match 97.4%; Score 114; DB 9; Length 23;
Best Local Similarity 95.7%; Pred. No. 5.3e-11;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCSARLVHQLA 23
DB 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 7

AAM50230 standard; protein; 205 AA.

AAM50230;

07-JAN-2002 (first entry)

Haemophilus influenzae protein HI 0638, involved in paracytosis.

ORF HI 0638; paracytosis; paracytin; lung; epithelium;

lower respiratory tract; infection; antimicrobial; drug screening;

therapy; vaccine.

Haemophilus influenzae.

EP1136557-A1.

26-SEP-2001.

21-MAR-2000; 2000EP-00201045.

21-MAR-2000; 2000EP-00201045.

(NEDERLANDEN MIN WELZIJN.

Van Den Schijfgaard, Van Alphen AJW;

WPI; 2001-640390/74.

N-PSDB; AAI70333.

New nucleic acids and polypeptides, which increases the permeability of

tissue or cell layer, useful as vaccine for treating lower respiratory

tract infections due to Haemophilus influenzae.

Claim 5; Page 19; 32pp; English.

The present sequence is that of protein HI 0638, which is encoded by

newly identified open reading frame (ORF) HI 0638 (see AAI70333) of

Haemophilus influenzae strain A960053, an isolate from a sputum sample of

a cystic fibrosis patient. The protein is involved in paracytosis. Clones

of E. coli DH5-alpha containing ORF HI 0636 (see AAI70332) and/or ORF HI

0638 showed significant increase in penetration of epithelial cell layers

of the human bronchial epithelial cell line NCI-H292. HI 0636 (see

AAM50230) and HI 0638 may be used to alter the permeability of tissues or

cell layers, especially lung epithelial cell layers. They will also be

useful for developing compounds that block paracytosis as well as to

obtain insight into the invasion mechanism used by various bacteria. This

may lead to novel compounds and/or approaches in the treatment of lower

respiratory tract infections due to H. influenzae. Vaccines containing or

based upon either the HI 0636 and/or HI 0638 protein or polypeptide, or a

microorganism expressing ORF HI 0636 and/or ORF HI 0638, are claimed

Sequence 205 AA;

Query Match 97.4%; Score 114; DB 4; Length 205;
Best Local Similarity 95.7%; Pred. No. 6.2e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVCSARLVHQLA 23
DB 3 NYHDIYALAGVCSARLVHQLA 25

RESULT 8

AEB08261 standard; protein; 205 AA.

AEB08261;

25-AUG-2005 (first entry)

Haemophilus influenzae ORF HI0638 protein, SEQ ID NO: 59.

Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;

antidiabetic; endocrine disease; gastrointestinal disease;

metabolic disorder; hormone deficiency; osteoporosis; osteopathic;

degeneration; musculoskeletal disease; ocular disease; ophthalmological;

neurodegenerative disease; neuroprotective; Alzheimer's disease;

neurotropic; neurological disease; Parkinson's disease; antiparkinsonian;

dementia; multiple sclerosis; immune disorder; Huntington's chorea;

anticonvulsant; genetic disorder; cardiovascular disease;

cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;

coronary artery disease; cardiac; vasotropic; obesity; anorectic;

nutritional disorder; vitamin deficiency; renal disease; nephrotropic;

genitourinary disease; hematological disease; antianemic; anemia;

autoimmune disease; immunosuppressive; immune deficiency;

immunostimulant; infectious disease; antimicrobial; infection;

erectile dysfunction; andrology; major depressive disorder;

antidepressant; psychiatric disorder; pain; analgesic;

bacterial infection; antibacterial; viral infection; virucide;

fungal infection; fungicide; parasitic infection; antiparasitic;

renal failure; antinephritis; antineumatic; cytostatic;

antimicrobial; hepatotropic; hepatitis B virus infection.

Haemophilus influenzae.

US2005136103-A1.

23-JUN-2005.

16-SEP-2004; 2004US-00942300.

17-SEP-2003; 2003US-00664989.

17-SEP-2003; 2003US-00665184.

17-SEP-2003; 2003US-0503615P.

(BENS/) BEN-SASSON S. A.

(COHE/) COHEN E.

Ben-Sasson SA, Cohen E;

WPI; 2005-444089/45.

Composition used for translocating effectors across barrier such as

epithelial cells during treatment of e.g. endocrine disorders comprises

effector sequentially coupled with counter ion and hydrophobic agent.

Disclosure; SEQ ID NO 59; 59pp; English.

The present invention relates to a pharmaceutical composition of

penetrating peptides for transmembrane delivery of effector. The

invention comprises the effector sequentially coupled with a counter ion

and at least one hydrophobic agent, where the effector is selectively

encapsulated into a complex. The invention is useful for translocating

effectors across a biological barrier such as epithelial cells and

endothelial cells during treatment and prevention of disease or

pathological conditions (including endocrine disorders, diabetes,

infertility, hormone deficiencies, osteoporosis, ophthalmological

disorders, neurodegenerative disorders, Alzheimer's disease,

Parkinson's disease, multiple sclerosis, Huntington's disease,

cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-

coagulable states, coronary disease, cerebrovascular events, metabolic

disorders, obesity, vitamin deficiencies, renal disorders, renal failure,

hematological disorders, anemia of different entities, immunologic and

neumatologic disorders, autoimmune diseases, immune deficiencies,

infectious diseases, viral infections, bacterial infections, fungal

infections, parasitic infections, neoplastic diseases, multi-factorial disorders, impotence, chronic pain, depression, different fibrosis states and short stature) and for mucosal vaccination against anthrax and hepatitis B. The present sequence is the Haemophilus influenzae ORF HI0638 protein containing penetrating peptide at N-terminal end.

Sequence 205 AA;

Query Match 97.4%; Score 114; DB 9; Length 205;
Best Local Similarity 95.7%; Pred. No. 6.2e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

NYHDIYVLAAGVCGSARLVHOLA 23
NYHDIYVLAAGVCGSARLVHOLA 25

RESULT 9
ADBI6888 standard; peptide; 23 AA.

ADBI6888;
20-NOV-2003 (first entry)

Escherichia coli YCFC penetrating peptide 3.

penetrating peptide; epithelial; endothelial; tight junction; diabetes;
infertility; hormone; vitamin deficiency; neurodegenerative;
cardiovascular; haematological; endocrine disorder; obesity;
neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
osteopathic; cyostatic; nootropic.

Escherichia coli.

WO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003WO-IB000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI; 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating

or preventing e.g. endocrine disorders.

Claim 2; Page 14; 60pp; English.

This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factor. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, haematological and endocrine disorders, as well as obesity and neoplastic disease. Accordingly, the peptides of this invention can be used in compositions that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic, cyostatic or nootropic activities. This peptide is from YCFC of Escherichia coli and is penetrating peptide 3 of the invention.

Sequence 23 AA;

Query Match 83.8%; Score 98; DB 6; Length 23;
Best Local Similarity 82.6%; Pred. No. 2e-08;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

NYHDIYVLAAGVCGSARLVHOLA 23
NYHDIYVLAAGVCGSARLVHOLA 23

RESULT 10
AEB08220 standard; peptide; 23 AA.

AEB08220;
25-AUG-2005 (first entry)

Escherichia coli YCFC penetrating peptide 3, SEQ ID NO: 3.

Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes;
antidiabetic; endocrine disease; gastrointestinal disease;
metabolic disorder; hormone deficiency; osteoporosis; osteopathic;
degeneration; musculoskeletal disease; ocular disease; ophthalmological;
neurodegenerative disease; neuroprotective; Alzheimer's disease;
nootropic; neurological disease; Parkinson's disease; antiparkinsonian;
dementia; multiple sclerosis; immune disorder; Huntington's chorea;
anticonvulsant; genetic disorder; cardiovascular disease;
cardiovascular-gen.; atherosclerosis; antiarteriosclerotic;
coronary artery disease; cardiac; vasotropic; obesity; anorectic;
nutritional disorder; vitamin deficiency; renal disease; nephrotropic;
genitourinary disease; hematological disease; antianemic; anemia;
autoimmune disease; immunosuppressive; immune deficiency;
immunostimulant; infectious disease; antimicrobial; infection;
erectile dysfunction; andrology; major depressive disorder;
antidepressant; psychiatric disorder; pain; analgesic;
bacterial infection; antibacterial; viral infection; vaccine;
fungal infection; fungicide; parasitic infection; antiparasitic;
renal failure; infertility; antineumatic; cyostatic;
antiinflammatory; hepatotropic; hepatitis B virus infection.

Escherichia coli.

US2005136103-A1.

23-JUN-2005.

16-SEP-2004; 2004US-00942300.

17-SEP-2003; 2003US-00664989.

17-SEP-2003; 2003US-00665184.

17-SEP-2003; 2003US-0503615P.

(BENS/) BEN-SASSON S A.

(COHE/) COHEN E.

Ben-Sasson SA, Cohen E;

WPI; 2005-444089/45.

Composition used for translocating effectors across barrier such as epithelial cells during treatment of e.g. endocrine disorders comprises effector sequentially coupled with counter ion and hydrophobic agent.

Claim 53; SEQ ID NO 3; 59pp; English.

The present invention relates to a pharmaceutical composition of penetrating peptides for trans epithelial delivery of effector. The invention comprises the effector sequentially coupled with a counter ion and at least one hydrophobic agent, where the effector is selectively encapsulated into a complex. The invention is useful for translocating

effectors across a biological barrier such as epithelial cells and endothelial cells during treatment and prevention of disease or pathological conditions (including endocrine disorders, diabetes, infertility, hormone deficiencies, osteoporosis, ophthalmological disorders, neurodegenerative disorders, Alzheimer's disease, dementia, Parkinson's disease, multiple sclerosis, Huntington's disease, hypocoagulable states, coronary disease, cerebrovascular events, metabolic disorders, obesity, vitamin deficiencies, renal disorders, renal failure, hematological disorders, anemia of different entities, immunologic and rheumatologic disorders, autoimmune diseases, immune deficiencies, infectious diseases, viral infections, bacterial infections, fungal infections, parasitic infections, neoplastic diseases, multi-factorial disorders, impotence, chronic pain, depression, different fibrosis states and short stature) and for mucosal vaccination against anthrax and hepatitis B. The present sequence is the Escherichia coli YCFC penetrating peptide. This sequence is used in the effective translocation of antimycoside antibiotics and antifungal agents across an epithelial barrier.

Sequence 23 AA;

Query Match 83.8%; Score 98; DB 9; Length 23;
Best Local Similarity 82.6%; Pred. No. 2e-08;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYTLALAGVQSARLVHQLA 23
DB 1 NYHDIYTLALAGVQSARLVHQLA 23

RESULT 11

ADB16908 standard; peptide; 24 AA.

ADB16908;

20-NOV-2003 (first entry)

Escherichia coli YCFC penetrating peptide 30.

penetrating peptide; epithelial; endothelial; tight junction; diabetes; infertility; hormone; vitamin deficiency; neurodegenerative; cardiovascular; haematological; endocrine disorder; obesity; neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic; osteopathic; cyostatic; nootropic.

Escherichia coli.

MO2003066859-A2.

14-AUG-2003.

07-FEB-2003; 2003MO-18000968.

07-FEB-2002; 2002US-0355396P.

(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.

Ben-Sasson SA, Cohen E;

WPI, 2003-697452/66.

New penetrating peptide, useful for preparing a composition for treating or preventing e.g. endocrine disorders.

Claim 2; Page 15; 60pp; English.

This invention relates to a novel peptide sequences capable of translocating across a biological barrier. Furthermore, it refers to methods that use these peptides to facilitate penetration of a biologically active effector molecule such as a drug or other therapeutic agent across biological barriers e.g. epithelial or endothelial cells

sealed by tight junctions. This peptide is derived from a bacterial toxin, an integral membrane or extracellular protein and can comprise an anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin or enzyme. The effector molecule, however, can comprise for example insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic factors. The penetrating peptide is useful for the treatment of various conditions including diabetes, infertility, hormone and vitamin deficiencies, neurodegenerative, cardiovascular, hematological and endocrine disorders, as well as obesity and neoplastic disease. Accordingly, the peptides of this invention can be used in compositions that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic, cyostatic or nootropic activities. This peptide is from YCFC of Escherichia coli and is penetrating peptide 30 of the invention.

Sequence 24 AA;

Query Match 83.8%; Score 98; DB 6; Length 24;
Best Local Similarity 82.6%; Pred. No. 2.1e-08;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYTLALAGVQSARLVHQLA 23
DB 2 NYHDIYTLALAGVQSARLVHQLA 24

RESULT 12

ARB08242 standard; peptide; 24 AA.

ARB08242;

25-AUG-2005 (first entry)

Escherichia coli YCFC penetrating peptide 30, SEQ ID NO: 25.

Pharmaceutical; therapeutic; endocrine disease; endocrine-gen.; diabetes; antidiabetic; endocrine disease; gastrointestinal disease; metabolic disorder; hormone deficiency; osteoporosis; osteopathic; degeneration; musculoskeletal disease; ocular disease; ophthalmological; neurodegenerative disease; neuroprotective; Alzheimer's disease; nootropic; neurological disease; Parkinson's disease; antiparkinsonian; dementia; multiple sclerosis; immune disorder; Huntingtons chorea; anticonvulsant; genetic disorder; cardiovascular disease; cardiovascular-gen.; atherosclerosis; antiarteriosclerotic; coronary artery disease; cardiac; vasotrophic; obesity; anorectic; nutritional disorder; vitamin deficiency; renal disease; nephrotropic; genitourinary disease; hematological disease; antiamebic; anemia; autoimmune disease; immunosuppressive; immune deficiency; immunostimulant; infectious disease; antimicrobial; infection; erectile dysfunction; andrology; major depressive disorder; antidepressant; psychiatric disorder; pain; analgesic; bacterial infection; antibacterial; viral infection; virucide; fungal infection; fungicide; parasitic infection; antiparasitic; renal failure; antifertility; antirheumatic; cyostatic; antiinflammatory; hepatotropic; hepatitis B virus infection.

Escherichia coli.

US2005136103-A1.

23-JUN-2005.

16-SEP-2004; 2004US-00942300.

17-SEP-2003; 2003US-00664989.

17-SEP-2003; 2003US-00665184.

17-SEP-2003; 2003US-0503615P.

(BENS/) BEN-SASSON S A.

(COHE/) COHEN E.

Ben-Sasson SA, Cohen E;

XX WPI; 2005-444089/45.
 DR Composition used for translocating effectors across barrier such as
 XX epithelial cells during treatment of e.g. endocrine disorders comprising
 PT effector sequentially coupled with counter ion and hydrophobic agent.
 XX
 XX Claim 46; SEQ ID NO 25; 59pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition of
 CC penetrating peptides for transendothelial delivery of effector. The
 CC invention comprises the effector sequentially coupled with a counter ion
 CC and at least one hydrophobic agent, where the effector is selectively
 CC encapsulated into a complex. The invention is useful for translocating
 CC effectors across a biological barrier such as epithelial cells and
 CC endothelial cells during treatment and prevention of disease or
 CC pathological conditions (including endocrine disorders, diabetes,
 CC infertility, hormone deficiencies, osteoporosis, ophthalmological
 CC disorders, neurodegenerative disorders, Alzheimer's disease, dementia,
 CC Parkinson's disease, multiple sclerosis, Huntington's disease, hypo-
 CC cardiovascular disorders, atherosclerosis, hyper-coagulable states, hypo-
 CC coagulable states, coronary disease, cerebrovascular events, metabolic
 CC disorders, obesity, vitamin deficiencies, renal disorders, renal failure,
 CC hematological disorders, anemia of different entities, immunologic and
 CC rheumatologic disorders, autoimmune diseases, immune deficiencies,
 CC infectious diseases, viral infections, bacterial infections, fungal
 CC infections, parasitic infections, neoplastic diseases, multi-factorial
 CC disorders, impotence, chronic pain, depression, different fibrosis states
 CC and short stature) and for mucosal vaccination against anthrax and
 CC hepatitis B. The present sequence is the Escherichia coli YFC
 CC penetrating peptide. This sequence is used in the effective translocation
 CC of aminoglycoside antibiotics and antifungal agents across an epithelial
 CC barrier.
 XX
 XX Sequence 24 AA;
 SQ
 Query Match 83.8%; Score 98; DB 9; Length 24;
 Best Local Similarity 82.6%; Pred. No. 2.1e-08;
 Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NYHDIYVLAVGVCOSARLVHQLA 23
 DB 2 NYDITLVLAVGICQSAARLVQQLA 24
 RESULT 13
 ADBI6922
 ID ADBI6922 standard; peptide; 25 AA.
 XX
 AC ADBI6922;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Escherichia coli peptide 3 linked to recombinant human insulin.
 XX
 KW recombinant human insulin; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cyostatic; nootropic; penetrating peptide; diabetes.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 XX Key Location/Qualifiers
 FH 1..23
 FT Peptide /note= "Penetrating peptide 3"
 FT Modified-site 25
 FT /note= "Recombinant human insulin peptide coupled to the
 FT C-terminus"
 XX
 XX MO200306859-A2.

PD 14-AUG-2003.
 XX
 XX 07-FEB-2003; 2003WO-IB000968.
 XX
 XX 07-FEB-2002; 2002US-0355396P.
 XX
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX
 XX Ben-Sasson SA, Cohen E;
 FL WPI; 2003-697452/66.
 DR
 XX
 XX WPI; 2003-697452/66.
 XX
 PT New penetrating peptide, useful for preparing a composition for treating
 PT or preventing e.g. endocrine disorders.
 XX
 XX Example 8; Page 42; 60pp; English.
 XX
 XX This invention relates to a novel peptide sequences capable of
 XX translocating across a biological barrier. Furthermore, it refers to
 XX methods that use these peptides to facilitate penetration of a
 XX biologically active effector molecule such as a drug or other therapeutic
 XX agent across biological barriers e.g. epithelial or endothelial cells
 XX sealed by tight junctions. This peptide is derived from a bacterial
 XX toxin, an integral membrane or extracellular protein and can comprise an
 XX anticoagulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
 XX or enzyme. The effector molecule, however, can comprise for example
 XX insulin, gonadotropin, erythropoietin, granulocyte/macocyte colony
 XX stimulating factor (GM-CSF), enkephalin, dalargin, or neurotrophic
 XX factors. The penetrating peptide is useful for the treatment of various
 XX conditions including diabetes, infertility, hormone and vitamin
 XX deficiencies, neurodegenerative, cardiovascular, haematological and
 XX endocrine disorders, as well as obesity and neoplastic disease.
 XX According to the invention, the peptides of this invention can be used in compositions
 XX that have neuroprotective, cardiac, antiarteriosclerotic, osteopathic,
 XX cyostatic or nootropic activities. This peptide sequence consists of the
 XX Escherichia coli penetrating peptide 3 coupled to recombinant human
 XX insulin in the absence of a detachable linker peptide as a cleavage site,
 XX used to deliver insulin across mucosal epithelia as a treatment for
 XX diabetes, in an exemplification of the invention.
 XX
 XX Sequence 25 AA;
 SQ
 Query Match 83.8%; Score 98; DB 6; Length 25;
 Best Local Similarity 82.6%; Pred. No. 2.2e-08;
 Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NYHDIYVLAVGVCOSARLVHQLA 23
 DB 1 NYDITLVLAVGICQSAARLVQQLA 23
 RESULT 14
 ADBI6906
 ID ADBI6906 standard; peptide; 25 AA.
 XX
 AC ADBI6906;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Escherichia coli peptide 3 linked to linearised insulin receptor peptide.
 XX
 KW linearised insulin receptor; epithelial; endothelial; tight junction;
 KW diabetes; infertility; hormone; vitamin deficiency; neurodegenerative;
 KW cardiovascular; haematological; endocrine disorder; obesity;
 KW neoplastic disease; neuroprotective; cardiac; antiarteriosclerotic;
 KW osteopathic; cyostatic; nootropic; penetrating peptide; diabetes.
 XX
 OS Synthetic.
 OS Escherichia coli.
 XX
 XX Key Location/Qualifiers
 FH 1..23
 FT Peptide /note= "Penetrating peptide 3"
 FT
 FT

```

FT Modified-site 25
FT /note= "Linearised insulin receptor peptide coupled to
FT the C-terminus"
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX WPI, 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Example 9; Page 43; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells
XX sealed by tight junctions. This peptide is derived from a bacterial
XX toxin, an integral membrane or extracellular protein and can comprise an
XX antioesulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX or enzyme. The effector molecule, however, can comprise for example
XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX stimulating factor (GM-CSF), enkephalin, dalaargin, or neurotrophic
XX factors. The penetrating peptide is useful for the treatment of various
XX conditions including diabetes, infertility, hormone and vitamin
XX deficiencies, neurodegenerative, cardiovascular, haematological and
XX endocrine disorders, as well as obesity and neoplastic disease.
XX Accordingly, the peptides of this invention can be used in compositions
XX that have neuroprotective, cardiant, antihypertensive, osteopathic,
XX cyostatic or neurotropic activities. This peptide sequence consists of the
XX Escherichia coli penetrating peptide 3 coupled to a linearised insulin
XX receptor, used to deliver insulin across mucosal epithelia as a treatment
XX for diabetes, in an exemplification of the invention.
XX
XX Sequence 25 AA;
XX
XX Query Match 83.8%; Score 98; DB 6; Length 25;
XX Best Local Similarity 82.6%; Pred. No. 2.2e-08;
XX Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 NYHDIYVATAGVCSARLVQOLA 23
XX ||:|||||:|||||:|||||
XX DB 1 NYDYITLALAGICQSARLVQOLA 23
XX
XX RESULT 15
XX ADB16904
XX ID ADB16904 standard; peptide; 26 AA.
XX
XX ADB16904;
XX
XX 20-NOV-2003 (first entry)
XX
XX E.coli penetrating peptide 3 linked to heparin without a cleavage site.
XX
XX heparin; epithelial; endothelial; tight junction; diabetes; infertility;
XX hormone; vitamin deficiency; neurodegenerative; cardiovascular;
XX haematological; endocrine disorder; obesity; neoplastic disease;
XX neuroprotective; cardiant; antihypertensive; osteopathic; cyostatic;
XX neurotropic; penetrating peptide.
XX
XX Synthetic.
XX
XX OS

```

```

XX Escherichia coli.
XX
XX Key Location/Qualifiers
XX Peptide 1..23
XX FT /note= "Penetrating peptide 3"
XX FT 26
XX FT Modified-site 25
XX FT /note= "Heparin peptide coupled to the C-terminus"
XX
XX WO2003066859-A2.
XX
XX 14-AUG-2003.
XX
XX 07-FEB-2003; 2003WO-IB000968.
XX
XX 07-FEB-2002; 2002US-0355396P.
XX
XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Ben-Sasson SA, Cohen E;
XX WPI, 2003-697452/66.
XX
XX New penetrating peptide, useful for preparing a composition for treating
XX or preventing e.g. endocrine disorders.
XX
XX Example 10; Page 44; 60pp; English.
XX
XX This invention relates to a novel peptide sequences capable of
XX translocating across a biological barrier. Furthermore, it refers to
XX methods that use these peptides to facilitate penetration of a
XX biologically active effector molecule such as a drug or other therapeutic
XX agent across biological barriers e.g. epithelial or endothelial cells
XX sealed by tight junctions. This peptide is derived from a bacterial
XX toxin, an integral membrane or extracellular protein and can comprise an
XX antioesulant, antibiotic, antipathogenic agent, immunomodulator, vitamin
XX or enzyme. The effector molecule, however, can comprise for example
XX insulin, gonadotropin, erythropoietin, granulocyte/monocyte colony
XX stimulating factor (GM-CSF), enkephalin, dalaargin, or neurotrophic
XX factors. The penetrating peptide is useful for the treatment of various
XX conditions including diabetes, infertility, hormone and vitamin
XX deficiencies, neurodegenerative, cardiovascular, haematological and
XX endocrine disorders, as well as obesity and neoplastic disease.
XX Accordingly, the peptides of this invention can be used in compositions
XX that have neuroprotective, cardiant, antihypertensive, osteopathic,
XX cyostatic or neurotropic activities. This peptide sequence consists of the
XX Escherichia coli penetrating peptide 3 coupled to heparin in the absence
XX of a detachable linker peptide as a cleavage site, used to deliver low
XX molecular weight heparin across mucosal epithelia in an exemplification
XX of the invention.
XX
XX Sequence 26 AA;
XX
XX Query Match 83.8%; Score 98; DB 6; Length 26;
XX Best Local Similarity 82.6%; Pred. No. 2.3e-08;
XX Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 NYHDIYVATAGVCSARLVQOLA 23
XX ||:|||||:|||||:|||||
XX DB 1 NYDYITLALAGICQSARLVQOLA 23
XX
XX Search completed: January 23, 2006, 11:29:00
XX Job time : 134 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 23, 2006, 12:09:04 ; Search time 62 Seconds
(without alignments)
155.001 Million cell updates/sec

Title: US-10-501-838A-29

Perfect score: 117

Sequence: 1 NYHDIYALAGVCSARLVHQLA 23

Scoring table: BLOSUM62

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Published Applications: AA_Main:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	100.0	23	4	US-10-665-184-29
2	117	100.0	23	5	US-10-942-300-29
3	117	100.0	23	5	US-10-501-838A-29
4	117	100.0	23	4	US-10-665-184-33
5	117	100.0	29	5	US-10-942-300-33
6	117	100.0	30	5	US-10-501-838A-33
7	114	97.4	23	4	US-10-665-184-1
8	114	97.4	23	5	US-10-942-300-1
9	114	97.4	23	5	US-10-501-838A-1
10	114	97.4	205	5	US-10-665-184-59
11	114	97.4	205	5	US-10-942-300-59
12	114	97.4	205	5	US-10-501-838A-53
13	98	83.8	23	4	US-10-665-184-3
14	98	83.8	23	5	US-10-942-300-3
15	98	83.8	23	5	US-10-501-838A-3
16	98	83.8	24	5	US-10-665-184-25
17	98	83.8	24	5	US-10-942-300-25
18	98	83.8	24	5	US-10-501-838A-25
19	98	83.8	25	5	US-10-501-838A-19
20	98	83.8	25	5	US-10-501-838A-23
21	98	83.8	26	5	US-10-501-838A-21
22	98	83.8	29	5	US-10-501-838A-18
23	98	83.8	30	5	US-10-665-184-22
24	98	83.8	30	5	US-10-942-300-22
25	98	83.8	30	5	US-10-501-838A-20
26	98	83.8	30	5	US-10-501-838A-22
27	98	83.8	31	4	US-10-665-184-35

28	98	83.8	31	5	US-10-942-300-35	Sequence 35, App1
29	98	83.8	31	5	US-10-501-838A-35	Sequence 35, App1
30	98	83.8	213	4	US-10-665-184-61	Sequence 61, App1
31	98	83.8	213	5	US-10-942-300-61	Sequence 61, App1
32	98	83.8	213	5	US-10-501-838A-55	Sequence 55, App1
33	98	83.8	213	5	US-10-665-184-26	Sequence 26, App1
34	98	83.8	23	5	US-10-942-300-26	Sequence 26, App1
35	98	83.8	23	5	US-10-501-838A-26	Sequence 26, App1
36	98	83.8	30	4	US-10-665-184-36	Sequence 36, App1
37	98	83.8	30	5	US-10-942-300-36	Sequence 36, App1
38	98	83.8	30	5	US-10-501-838A-36	Sequence 36, App1
39	98	83.8	23	4	US-10-665-184-2	Sequence 2, App1
40	98	83.8	23	5	US-10-942-300-2	Sequence 2, App1
41	98	83.8	23	5	US-10-501-838A-2	Sequence 2, App1
42	98	83.8	203	4	US-10-665-184-60	Sequence 60, App1
43	98	83.8	203	5	US-10-942-300-60	Sequence 60, App1
44	98	83.8	203	5	US-10-501-838A-54	Sequence 54, App1
45	98	83.8	22	4	US-10-665-184-27	Sequence 27, App1

ALIGNMENTS

RESULT 1
US-10-665-184-29
; Sequence 29, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across ;
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent version 3.2
; SEQ ID NO 29
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-665-184-29
Query Match 100.0%; Score 117; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 NYHDIYALAGVCSARLVHQLA 23
QY 1 NYHDIYALAGVCSARLVHQLA 23
US-10-942-300-29
; Sequence 29, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615

```

; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-942-300-29

Query Match          100.0%; Score 117; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYHDIYALAGVCSARLVHQLA 23
Db 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 3
US-10-501-838A-29
; Sequence 29, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
; CURRENT APPLICATION NUMBER: US/10/501,838A
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 29
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-501-838A-29

Query Match          100.0%; Score 117; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYHDIYALAGVCSARLVHQLA 23
Db 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 4
US-10-665-184-33
; Sequence 33, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial
```

```

; FEATURE:
; OTHER INFORMATION: Penetrating peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: ACETYLATION
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (29)..(29)
; OTHER INFORMATION: wherein Xaa is Lysine-NH2
US-10-665-184-33

Query Match          100.0%; Score 117; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 6.7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYHDIYALAGVCSARLVHQLA 23
Db 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 5
US-10-942-300-33
; Sequence 33, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; PRIOR FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; PRIOR FILING DATE: 2003-09-17
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial
; OTHER INFORMATION: Acylated Penetrating Peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (29)..(29)
; OTHER INFORMATION: wherein Xaa is Lysine having a free amino group that is acylated
US-10-942-300-33

Query Match          100.0%; Score 117; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 6.7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 NYHDIYALAGVCSARLVHQLA 23
Db 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 6
US-10-501-838A-33
; Sequence 33, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
```


CURRENT APPLICATION NUMBER: US/10/501,838A
CURRENT FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: US 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 33
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic: penetrating peptide
NAME/KEY: MISC FEATURE
LOCATION: (27)..(27)
OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
FEATURE:
OTHER INFORMATION: groups of the lysine residue
NAME/KEY: MISC FEATURE
LOCATION: (30)..(30)
OTHER INFORMATION: wherein the penetrating peptide is acylated via the free amino
FEATURE:
OTHER INFORMATION: groups of the lysine residue
NAME/KEY: MISC FEATURE
LOCATION: (30)..(30)
OTHER INFORMATION: wherein another molecule can be coupled to the penetrating
OTHER INFORMATION: peptide via the free amino groups of the lysine residue
US-10-501-838A-33

Query Match 100.0%; Score 117; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 7e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSARLVHQLA 23
Db 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 7
US-10-665-184-1
Sequence 1, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
APPLICANT: Cohen, Binat
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-501CIP
CURRENT APPLICATION NUMBER: US/10/665,184
CURRENT FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 23
TYPE: PRT
ORGANISM: haemophilus influenzae
US-10-665-184-1

Query Match 97.4%; Score 114; DB 4; Length 23;
Best Local Similarity 95.7%; Pred. No. 1.6e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSARLVHQLA 23
Db 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 8
US-10-942-300-1
Sequence 1, Application US/10942300
Publication No. US20050136103A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel
APPLICANT: Cohen, Binat
TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
FILE REFERENCE: 24348-503
CURRENT APPLICATION NUMBER: US/10/942,300
CURRENT FILING DATE: 2004-09-16
PRIOR APPLICATION NUMBER: 10/665,184
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 10/664,989
PRIOR FILING DATE: 2003-09-17
PRIOR APPLICATION NUMBER: 60/503,615
PRIOR FILING DATE: 2003-09-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 23
TYPE: PRT
ORGANISM: haemophilus influenzae
US-10-942-300-1

Query Match 97.4%; Score 114; DB 5; Length 23;
Best Local Similarity 95.7%; Pred. No. 1.6e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSARLVHQLA 23
Db 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 9
US-10-501-838A-1
Sequence 1, Application US/10501838A
Publication No. US20050215478A1
GENERAL INFORMATION:
APPLICANT: Ben-Sasson, Shmuel A.
APPLICANT: Cohen, Binat
TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across
FILE REFERENCE: 24348-501 NATL
CURRENT APPLICATION NUMBER: US/10/501,838A
CURRENT FILING DATE: 2004-07-19
PRIOR APPLICATION NUMBER: PCT/IB03/00968
PRIOR FILING DATE: 2003-02-07
PRIOR APPLICATION NUMBER: 60/355,396
PRIOR FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 72
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 23
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-10-501-838A-1

Query Match 97.4%; Score 114; DB 5; Length 23;
Best Local Similarity 95.7%; Pred. No. 1.6e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYHDIYALAGVCSARLVHQLA 23
Db 1 NYHDIYALAGVCSARLVHQLA 23

RESULT 10
US-10-665-184-59
Sequence 59, Application US/10665184
Publication No. US20040146549A1
GENERAL INFORMATION:

```

; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-665-184-59

Query Match          97.4%; Score 114; DB 4; Length 205;
Best Local Similarity 95.7%; Pred. No. 1.5e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NYHDIVLALAGVCSARLVHQLA 23
Db 3 NYHDIVLALAGVCSARLVHQLA 25

RESULT 11
US-10-942-300-59
; Sequence 59, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 59
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-942-300-59

Query Match          97.4%; Score 114; DB 5; Length 205;
Best Local Similarity 95.7%; Pred. No. 1.5e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NYHDIVLALAGVCSARLVHQLA 23
Db 3 NYHDIVLALAGVCSARLVHQLA 25

RESULT 12
US-10-501-838A-53
; Sequence 53, Application US/10501838A
; Publication No. US20050215478A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel A.
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501 NATL
```

```

; CURRENT APPLICATION NUMBER: US/10/501,838A
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 53
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-501-838A-53

Query Match          97.4%; Score 114; DB 5; Length 205;
Best Local Similarity 95.7%; Pred. No. 1.5e-09;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NYHDIVLALAGVCSARLVHQLA 23
Db 3 NYHDIVLALAGVCSARLVHQLA 25

RESULT 13
US-10-665-184-3
; Sequence 3, Application US/10665184
; Publication No. US20040146549A1
; GENERAL INFORMATION:
; APPLICANT: Ben-Sasson, Shmuel
; APPLICANT: Cohen, Elina
; TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-501CIP
; CURRENT APPLICATION NUMBER: US/10/665,184
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/IB03/00968
; PRIOR FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: 60/355,396
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-665-184-3

Query Match          83.8%; Score 98; DB 4; Length 23;
Best Local Similarity 82.6%; Pred. No. 4.9e-08;
Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NYHDIVLALAGVCSARLVHQLA 23
Db 1 NYDITLALAGICQARLVHQLA 23

RESULT 14
US-10-942-300-3
; Sequence 3, Application US/10942300
; Publication No. US20050136103A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Elina
; APPLICANT: Ben-Sasson, Shmuel
; TITLE OF INVENTION: Compositions Capable of Facilitating Penetration Across a
; FILE REFERENCE: 24348-503
; CURRENT APPLICATION NUMBER: US/10/942,300
; CURRENT FILING DATE: 2004-09-16
; PRIOR APPLICATION NUMBER: 10/665,184
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 10/664,989
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: 60/503,615
```

/ PRIOR FILING DATE: 2003-09-17
 / NUMBER OF SEQ ID NOS: 65
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 3
 / LENGTH: 23
 / TYPE: PRT
 / ORGANISM: Escherichia coli
 US-10-942-300-3

Query Match 83.8%; Score 98; DB 5; Length 23;
 Best Local Similarity 82.6%; Pred. No. 4.9e-08;
 Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVQGSARLVHQLA 23
 ||:|||||:|||||
 DB 1 NYDITLALAGICQGSARLVQQLA 23

RESULT 15
 US-10-501-838A-3
 / Sequence 3, Application US/10501838A
 / Publication No. US20050215478A1
 / GENERAL INFORMATION:
 / APPLICANT: Ben-Sasson, Shmuel A.
 / APPLICANT: Cohen, Eilat
 / TITLE OF INVENTION: Amino Acid Sequences Capable of Facilitating Penetration Across a
 / TITLE OF INVENTION: Biological Barrier
 / FILE REFERENCE: 24348-501 NATL
 / CURRENT APPLICATION NUMBER: US/10/501, 838A
 / CURRENT FILING DATE: 2004-07-19
 / PRIOR APPLICATION NUMBER: PCT/IB03/00968
 / PRIOR FILING DATE: 2003-02-07
 / PRIOR APPLICATION NUMBER: US 60/355,396
 / PRIOR FILING DATE: 2002-02-07
 / NUMBER OF SEQ ID NOS: 72
 / SOFTWARE: PatentIn version 3.2
 / SEQ ID NO 3
 / LENGTH: 23
 / TYPE: PRT
 / ORGANISM: Escherichia coli
 US-10-501-838A-3

Query Match 83.8%; Score 98; DB 5; Length 23;
 Best Local Similarity 82.6%; Pred. No. 4.9e-08;
 Matches 19; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYHDIYALAGVQGSARLVHQLA 23
 ||:|||||:|||||
 DB 1 NYDITLALAGICQGSARLVQQLA 23

Search completed: January 23, 2006, 12:13:03
 Job time : 62 secs

This Page Blank (upside)

This Page Blank (upside)